GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 02:28:24 ; Search time 42.4286 Seconds
(without alignments)
39.956 Million cell updates/sec

Title: US-10-029-926B-8
Perfect score: 29
Sequence: 1 MRAPVI 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	6	2	AAW77197	Aaw77197 Pharmaceu
2	29	100.0	6	5	ABG78133	Abg78133 Human Fv
3	29	100.0	6	5	ABG91824	Abg91824 Human ant
4	29	100.0	246	5	ABG78329	Abg78329 Human Fv
5	29	100.0	246	5	ABG92026	Abg92026 Antibody
6	29	100.0	256	5	ABG78334	Abg78334 Human Fv
7	29	100.0	256	5	ABG92025	Abg92025 Antibody
8	29	100.0	266	5	ABG92020	Abg92020 Human ant
9	29	100.0	277	5	ABG78150	Abg78150 Human Fv
10	29	100.0	277	5	ABG91841	Abg91841 Human ant
11	29	100.0	464	5	ABG78151	Abg78151 Human Fv
12	29	100.0	464	5	ABG92021	Abg92021 Antibody
13	29	100.0	464	5	ABG91842	Abg91842 Human ant
14	28	96.6	133	4	AAU42688	Aau42688 Propionib
15	28	96.6	133	6	ABM39207	Abm39207 Propionib
16	28	96.6	438	5	AAE14569	Aae14569 Human rho
17	28	96.6	438	7	ADA13449	Ada13449 Human rho
18	28	96.6	438	7	ADA13415	Ada13415 Human rho
19	28	96.6	470	7	ADA13459	Ada13459 Mouse rho
20	28	96.6	541	4	AAG90242	Aag90242 C glutami
21	27	93.1	116	4	AAU44808	Aau44808 Propionib
22	27	93.1	116	6	ABM41327	Abm41327 Propionib
23	26	89.7	33	4	AAM82945	Aam82945 Human imm
24	26	89.7	58	4	AAU56719	Aau56719 Propionib
25	26	89.7	58	6	ABM53238	Abm53238 Propionib
26	26	89.7	226	5	ABP73581	Abp73581 Candida a
27	26	89.7	286	4	AAG98915	Aag98915 E. coli g
28	26	89.7	286	6	ABU15292	Abu15292 Protein e
29	26	89.7	290	6	ABU28269	Abu28269 Protein e
30	26	89.7	350	2	AAY35225	Aay35225 Chlamydia
31	26	89.7	405	7	ADC87283	Adc87283 Human GPC
32	26	89.7	1449	4	ABB66060	Abb66060 Drosophil
33	25	86.2	21	2	AAY12754	Aay12754 Human 5'
34	25	86.2	82	5	ABP08294	Abp08294 Human ORF
35	25	86.2	86	3	AAG18780	Aag18780 Zea mays
36	25	86.2	86	4	AAM90282	Aam90282 Human imm
37	25	86.2	89	3	AAG26861	Aag26861 Zea mays
38	25	86.2	102	6	ABM65637	Abm65637 Propionib
39	25	86.2	116	5	ABP32292	Abp32292 Human ORF
40	25	86.2	120	7	ADC97061	Adc97061 E. faeciu
41	25	86.2	125	2	AAY11311	Aay11311 S. pneumo
42	25	86.2	134	4	AAU45601	Aau45601 Propionib
43	25	86.2	134	6	ABM42120	Abm42120 Propionib
44	25	86.2	158	4	AAU42587	Aau42587 Propionib
45	25	86.2	158	6	ABM39106	Abm39106 Propionib

ALIGNMENTS

RESULT 1
AAW77197
ID AAW77197 standard; peptide; 6 AA.
XX
AC AAW77197;
XX
DT 23-NOV-1998 (first entry)
XX
DE Pharmaceutically active peptide 38.
XX
KW Pharmacuetically active peptide; target; organ; lymphocyte; treatment;
KW pharmaceutical agent; disease; radioactive isotope; imaging agent.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9839469-A1.
XX
PD 11-SEP-1998.
XX
PF 04-MAR-1998; 98WO-US004188.
XX
PR 04-MAR-1997; 97US-0039509P.
PR 04-MAR-1997; 97US-00810074.
XX
PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
XX
PI Panet A, Hagai Y, Lazarovits J, Nimrod A, Vogel T, Levanon A;
PI Zeelon E, Belkind A, Golan I;
XX
DR WPI; 1998-495863/42.
XX
PT New peptide(s) binding targets in organs and lymphocytes - used for the
PT targetted delivery of toxins, anti-cancer drugs and cardiovascular agents
PT to arteries, veins, placenta, liver.
XX
PS Claim 76; Page 99; 114pp; English.
XX
CC Sequences shown in AAW77160 to AAW77214 and AAW79167 represent non-
CC naturally- occurring pharmacuetically active peptides. These novel
CC peptides specifically bind to undetermined and determined targets in
CC various organs and in lymphocytes. The peptides can be used in
CC compositions, where they can be linked to pharmaceutical agents, to treat
CC various diseases and conditions. The peptides or chimeric polypeptides
CC comprising these pharmaceutically active peptides and a second peptide
CC may be labelled with a marker (radioactive isotope, etc) to form an
CC imaging agent. This agent is used to bind an organ so that the organ can
CC be imaged

XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 1 MRAPVI 6

RESULT 2
ABG78133
ID ABG78133 standard; peptide; 6 AA.
XX
AC ABG78133;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human Fv molecule hypervariable region related peptide #8.
XX
KW Human; Fv molecule; hypervariable region; single chain Fv; cytostatic;
KW disulfide Fv; dsFv; scFv; cancer; carcinoma; sarcoma; leukaemia; adenoma;
KW lymphoma; myeloma; blastoma; seminoma; melanoma; acute myeloid leukaemia.
XX
OS Homo sapiens.
XX
PN WO200259264-A2.
XX
PD 01-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049440.
XX
PR 29-DEC-2000; 2000US-00751181.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Hagai Y, Lazarovits J, Guy R, Lipschitz O, Szanton E, Levanon A;
PI Plaksin D, Peretz T;
XX
DR WPI; 2002-619166/66.
XX
PT Novel peptide/polypeptide for cancer therapy has Fv molecule, construct
PT or fragment, or construct of fragment with enhanced binding
PT characteristics so as to selectively bind target cell in favor of other
PT cells.
XX
PS Claim 2; Page 76; 232pp; English.
XX
CC The invention relates to a peptide or polypeptide comprising an Fv
CC molecule, a construct or fragments or a construct of a fragment with
CC enhanced binding characteristics which selectively and/or specifically
CC binds to a target cell in favour of other cells, where binding is
CC primarily determined by a first hypervariable region and Fv is a single
CC chain Fv (scFv) or a disulfide Fv (dsFv). The peptide, optionally in
CC association with or attached, coupled, combined, linked or fused to a
CC pharmaceutical agent, is useful in the manufacture of a medicament, where
CC the medicament has activity against a diseased cell, preferably a cancer
CC cell (selected from carcinoma, sarcoma, leukaemia, adenoma, lymphoma,
CC myeloma, blastoma, seminoma, and melanoma, where the leukaemia cell is an
CC acute myeloid leukaemia cell). The peptide is also useful for preparing a
CC composition for use in inhibiting the growth of a diseased or cancer
CC cell. This sequence represents a human Fv molecule hypervariable region
CC related peptide of the invention
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 1 MRAPVI 6

RESULT 3
ABG91824
ID ABG91824 standard; peptide; 6 AA.
XX
AC ABG91824;
XX
DT 04-DEC-2002 (first entry)
XX
DE Human antibody fragment #8.
XX
KW Human; antibody; epitope; cancer; tumour; cell rolling; inflammation;
KW metastasis; hypervariable region; autoimmune disease; thrombosis;
KW restenosis; leukaemia; inflammatory disease; cardiovascular disease;
KW myocardial infarction; retinopathic disease; abnormal platelet function;
KW sulphated tyrosine-dependent protein-protein interaction.
XX
OS Homo sapiens.
XX
PN WO200253700-A2.
XX
PD 11-JUL-2002.
XX
PF 31-DEC-2001; 2001WO-US049442.
XX
PR 29-DEC-2000; 2000US-00751181.
PR 29-DEC-2000; 2000US-0258948P.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Lazarovits J, Hagai Y, Plaksin D, Vogel T, Nimrod A, Mar-Haim H;
PI Szanthon E, Richter T, Amit B, Kooperman L, Peretz T, Levanon A;
XX
DR WPI; 2002-674776/72.
XX
PT Novel isolated epitope present on cancer cells and important in
PT physiological phenomena such as cell rolling, metastasis and
PT inflammation, for treating autoimmune, inflammatory or cardiovascular
PT diseases, and cancer.
XX
PS Claim 1; Page 228; 0pp; English.
XX
CC The invention relates to an isolated epitope present on cancer cells and
CC important in physiological phenomena such as cell rolling, metastasis and
CC inflammation, where the epitope is capable of being bound by an antibody,
CC its antigen-binding fragment or its complex comprising at least one
CC antibody or its binding fragment having a first hypervariable region. The
CC epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
CC disease, thrombosis, restenosis, metastasis, growth and/or replication of
CC tumour or leukaemia cells, increase in number of tumour or leukaemia
CC cells in a patient, cell-cell, cell-matrix, platelet-matrix, platelet-
CC platelet and/or cell-platelet adhesion or aggregation, for increasing
CC mortality of tumour or leukaemia cells, for increasing the susceptibility
CC of diseased cells to damage by anti-disease, anti-cancer or anti-
CC leukaemia agents, or for decreasing the number of tumour or leukaemia
CC cells in a patient, or in the manufacture of a medicament for the above
CC mentioned purposes. The epitopes are useful for diagnosing and treating
CC diseases such as cancer, leukaemia, autoimmune diseases, inflammatory
CC diseases, cardiovascular diseases such as myocardial infarction,
CC retinopathic diseases and other diseases mediated by abnormal platelet
CC function and diseases caused by sulphated tyrosine-dependent protein-
CC protein interactions. This sequence represents a human antibody fragment
CC of the invention
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRAPVI 6
              ||||||
Db          1 MRAPVI 6


RESULT 4
ABG78329
ID   ABG78329 standard; protein; 246 AA.
XX
AC   ABG78329;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Human Fv molecule hypervariable region related peptide #204.
XX
KW   Human; Fv molecule; hypervariable region; single chain Fv; cytostatic;
KW   disulfide Fv; dsFv; scFv; cancer; carcinoma; sarcoma; leukaemia; adenoma;
KW   lymphoma; myeloma; blastoma; seminoma; melanoma; acute myeloid leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200259264-A2.
XX
PD   01-AUG-2002.
XX
PF   31-DEC-2001; 2001WO-US049440.
XX
PR   29-DEC-2000; 2000US-00751181.
XX
PA   (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI   Hagai Y,  Lazarovits J,  Guy R,  Lipschitz O,  Szanton E,  Levanon A;
PI   Plaksin D,  Peretz T;
XX
DR   WPI; 2002-619166/66.
XX
PT   Novel peptide/polypeptide for cancer therapy has Fv molecule, construct
PT   or fragment, or construct of fragment with enhanced binding
PT   characteristics so as to selectively bind target cell in favor of other
PT   cells.
XX
PS   Disclosure; Page 44-45; 232pp; English.
XX
CC   The invention relates to a peptide or polypeptide comprising an Fv
CC   molecule, a construct or fragments or a construct of a fragment with
CC   enhanced binding characteristics which selectively and/or specifically
CC   binds to a target cell in favour of other cells, where binding is
CC   primarily determined by a first hypervariable region and Fv is a single
CC   chain Fv (scFv) or a disulfide Fv (dsFv). The peptide, optionally in
CC   association with or attached, coupled, combined, linked or fused to a
CC   pharmaceutical agent, is useful in the manufacture of a medicament, where
CC   the medicament has activity against a diseased cell, preferably a cancer
CC   cell (selected from carcinoma, sarcoma, leukaemia, adenoma, lymphoma,
CC   myeloma, blastoma, seminoma, and melanoma, where the leukaemia cell is an
CC   acute myeloid leukaemia cell). The peptide is also useful for preparing a
CC   composition for use in inhibiting the growth of a diseased or cancer
CC   cell. This sequence represents a human Fv molecule hypervariable region
CC   related peptide of the invention
XX
SQ   Sequence 246 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db        100 MRAPVI 105


RESULT 5
ABG92026
ID   ABG92026 standard; protein; 246 AA.
XX
AC   ABG92026;
XX
DT   04-DEC-2002  (first entry)
XX
DE   Antibody protein #5.
XX
KW   Antibody; epitope; cancer; tumour; cell rolling; inflammation;
KW   metastasis; hypervariable region; autoimmune disease; thrombosis;
KW   restenosis; leukaemia; inflammatory disease; cardiovascular disease;
KW   myocardial infarction; retinopathic disease; abnormal platelet function;
KW   sulphated tyrosine-dependent protein-protein interaction.
XX
OS   Unidentified.
XX
PN   WO200253700-A2.
XX
PD   11-JUL-2002.
XX
PF   31-DEC-2001; 2001WO-US049442.
XX
PR   29-DEC-2000; 2000US-00751181.
PR   29-DEC-2000; 2000US-0258948P.
XX
PA   (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI   Lazarovits J,  Hagai Y,  Plaksin D,  Vogel T,  Nimrod A,  Mar-Haim H;
PI   Szanthon E,  Richter T,  Amit B,  Kooperman L,  Peretz T,  Levanon A;
XX
DR   WPI; 2002-674776/72.
XX
PT   Novel isolated epitope present on cancer cells and important in
PT   physiological phenomena such as cell rolling, metastasis and
PT   inflammation, for treating autoimmune, inflammatory or cardiovascular
PT   diseases, and cancer.
XX
PS   Disclosure; Fig 52; 0pp; English.
XX
CC   The invention relates to an isolated epitope present on cancer cells and
CC   important in physiological phenomena such as cell rolling, metastasis and
CC   inflammation, where the epitope is capable of being bound by an antibody,
CC   its antigen-binding fragment or its complex comprising at least one
CC   antibody or its binding fragment having a first hypervariable region. The
CC   epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
CC   disease, thrombosis, restenosis, metastasis, growth and/or replication of
CC   tumour or leukaemia cells, increase in number of tumour or leukaemia
CC   cells in a patient, cell-cell, cell-matrix, platelet-matrix, platelet-
CC   platelet and/or cell-platelet adhesion or aggregation, for increasing
CC   mortality of tumour or leukaemia cells, for increasing the susceptibility
CC   of diseased cells to damage by anti-disease, anti-cancer or anti-
CC   leukaemia agents, or for decreasing the number of tumour or leukaemia
CC   cells in a patient, or in the manufacture of a medicament for the above
CC   mentioned purposes. The epitopes are useful for diagnosing and treating
CC   diseases such as cancer, leukaemia, autoimmune diseases, inflammatory
CC   diseases, cardiovascular diseases such as myocardial infarction,
CC   retinopathic diseases and other diseases mediated by abnormal platelet
CC   function and diseases caused by sulphated tyrosine-dependent protein-
CC   protein interactions. This sequence represents an antibody protein of the
CC   invention
XX
SQ   Sequence 246 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db        100 MRAPVI 105


RESULT 6
ABG78334
```

ID ABG78334 standard; protein; 256 AA.
XX
AC ABG78334;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human Fv molecule hypervariable region related peptide #209.
XX
KW Human; Fv molecule; hypervariable region; single chain Fv; cytostatic;
KW disulfide Fv; dsFv; scFv; cancer; carcinoma; sarcoma; leukaemia; adenoma;
KW lymphoma; myeloma; blastoma; seminoma; melanoma; acute myeloid leukaemia.
XX
OS Homo sapiens.
XX
PN WO200259264-A2.
XX
PD 01-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049440.
XX
PR 29-DEC-2000; 2000US-00751181.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Hagai Y, Lazarovits J, Guy R, Lipschitz O, Szanton E, Levanon A;
PI Plaksin D, Peretz T;
XX
DR WPI; 2002-619166/66.
XX
PT Novel peptide/polypeptide for cancer therapy has Fv molecule, construct
PT or fragment, or construct of fragment with enhanced binding
PT characteristics so as to selectively bind target cell in favor of other
PT cells.
XX
PS Example 9; Page 90; 232pp; English.
XX
CC The invention relates to a peptide or polypeptide comprising an Fv
CC molecule, a construct or fragments or a construct of a fragment with
CC enhanced binding characteristics which selectively and/or specifically
CC binds to a target cell in favour of other cells, where binding is
CC primarily determined by a first hypervariable region and Fv is a single
CC chain Fv (scFv) or a disulfide Fv (dsFv). The peptide, optionally in
CC association with or attached, coupled, combined, linked or fused to a
CC pharmaceutical agent, is useful in the manufacture of a medicament, where
CC the medicament has activity against a diseased cell, preferably a cancer
CC cell (selected from carcinoma, sarcoma, leukaemia, adenoma, lymphoma,
CC myeloma, blastoma, seminoma, and melanoma, where the leukaemia cell is an
CC acute myeloid leukaemia cell). The peptide is also useful for preparing a
CC composition for use in inhibiting the growth of a diseased or cancer
CC cell. This sequence represents a human Fv molecule hypervariable region
CC related peptide of the invention
XX
SQ Sequence 256 AA;

Query Match 100.0%; Score 29; DB 5; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 100 MRAPVI 105

RESULT 7
ABG92025
ID ABG92025 standard; protein; 256 AA.
XX
AC ABG92025;
XX
DT 04-DEC-2002 (first entry)
XX
DE Antibody biotag #1.
XX
KW Antibody; antibody biotag; cancer; tumour; cell rolling; inflammation;
KW metastasis; hypervariable region; autoimmune disease; thrombosis;
KW restenosis; leukaemia; inflammatory disease; cardiovascular disease;
KW myocardial infarction; retinopathic disease; abnormal platelet function;
KW sulphated tyrosine-dependent protein-protein interaction.
XX
OS Unidentified.
XX
PN WO200253700-A2.
XX
PD 11-JUL-2002.
XX
PF 31-DEC-2001; 2001WO-US049442.
XX
PR 29-DEC-2000; 2000US-00751181.
PR 29-DEC-2000; 2000US-0258948P.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Lazarovits J, Hagai Y, Plaksin D, Vogel T, Nimrod A, Mar-Haim H;
PI Szanthon E, Richter T, Amit B, Kooperman L, Peretz T, Levanon A;
XX
DR WPI; 2002-674776/72.
XX
PT Novel isolated epitope present on cancer cells and important in
PT physiological phenomena such as cell rolling, metastasis and
PT inflammation, for treating autoimmune, inflammatory or cardiovascular
PT diseases, and cancer.
XX
PS Disclosure; Fig 51; 0pp; English.
XX
CC The invention relates to an isolated epitope present on cancer cells and
CC important in physiological phenomena such as cell rolling, metastasis and
CC inflammation, where the epitope is capable of being bound by an antibody,
CC its antigen-binding fragment or its complex comprising at least one
CC antibody or its binding fragment having a first hypervariable region. The
CC epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
CC disease, thrombosis, restenosis, metastasis, growth and/or replication of
CC tumour or leukaemia cells, increase in number of tumour or leukaemia
CC cells in a patient, cell-cell, cell-matrix, platelet-matrix, platelet-
CC platelet and/or cell-platelet adhesion or aggregation, for increasing
CC mortality of tumour or leukaemia cells, for increasing the susceptibility
CC of diseased cells to damage by anti-disease, anti-cancer or anti-
CC leukaemia agents, or for decreasing the number of tumour or leukaemia
CC cells in a patient, or in the manufacture of a medicament for the above
CC mentioned purposes. The epitopes are useful for diagnosing and treating
CC diseases such as cancer, leukaemia, autoimmune diseases, inflammatory
CC diseases, cardiovascular diseases such as myocardial infarction,
CC retinopathic diseases and other diseases mediated by abnormal platelet
CC function and diseases caused by sulphated tyrosine-dependent protein-
CC protein interactions. This sequence represents an antibody biotag used in
CC the scope of the invention
XX
SQ Sequence 256 AA;

Query Match 100.0%; Score 29; DB 5; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 100 MRAPVI 105

RESULT 8
ABG92020
ID ABG92020 standard; protein; 266 AA.
XX
AC ABG92020;
XX
DT 04-DEC-2002 (first entry)
XX
DE Human antibody fragment #204.

XX
KW Human; antibody; epitope; cancer; tumour; cell rolling; inflammation;
KW metastasis; hypervariable region; autoimmune disease; thrombosis;
KW restenosis; leukaemia; inflammatory disease; cardiovascular disease;
KW myocardial infarction; retinopathic disease; abnormal platelet function;
KW sulphated tyrosine-dependent protein-protein interaction.
XX
OS Homo sapiens.
XX
PN WO200253700-A2.
XX
PD 11-JUL-2002.
XX
PF 31-DEC-2001; 2001WO-US049442.
XX
PR 29-DEC-2000; 2000US-00751181.
PR 29-DEC-2000; 2000US-0258948P.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Lazarovits J, Hagai Y, Plaksin D, Vogel T, Nimrod A, Mar-Haim H;
PI Szanthon E, Richter T, Amit B, Kooperman L, Peretz T, Levanon A;
XX
DR WPI; 2002-674776/72.
XX
PT Novel isolated epitope present on cancer cells and important in
PT physiological phenomena such as cell rolling, metastasis and
PT inflammation, for treating autoimmune, inflammatory or cardiovascular
PT diseases, and cancer.
XX
PS Disclosure; Page 309-310; 0pp; English.
XX
CC The invention relates to an isolated epitope present on cancer cells and
CC important in physiological phenomena such as cell rolling, metastasis and
CC inflammation, where the epitope is capable of being bound by an antibody,
CC its antigen-binding fragment or its complex comprising at least one
CC antibody or its binding fragment having a first hypervariable region. The
CC epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
CC disease, thrombosis, restenosis, metastasis, growth and/or replication of
CC tumour or leukaemia cells, increase in number of tumour or leukaemia
CC cells in a patient, cell-cell, cell-matrix, platelet-matrix, platelet-
CC platelet and/or cell-platelet adhesion or aggregation, for increasing
CC mortality of tumour or leukaemia cells, for increasing the susceptibility
CC of diseased cells to damage by anti-disease, anti-cancer or anti-
CC leukaemia agents, or for decreasing the number of tumour or leukaemia
CC cells in a patient, or in the manufacture of a medicament for the above
CC mentioned purposes. The epitopes are useful for diagnosing and treating
CC diseases such as cancer, leukaemia, autoimmune diseases, inflammatory
CC diseases, cardiovascular diseases such as myocardial infarction,
CC retinopathic diseases and other diseases mediated by abnormal platelet
CC function and diseases caused by sulphated tyrosine-dependent protein-
CC protein interactions. This sequence represents a human antibody fragment
CC of the invention
XX
SQ Sequence 266 AA;

Query Match 100.0%; Score 29; DB 5; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 121 MRAPVI 126

RESULT 9
ABG78150
ID ABG78150 standard; protein; 277 AA.
XX
AC ABG78150;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human Fv molecule hypervariable region related peptide #25.
XX
KW Human; Fv molecule; hypervariable region; single chain Fv; cytostatic;
KW disulfide Fv; dsFv; scFv; cancer; carcinoma; sarcoma; leukaemia; adenoma;
KW lymphoma; myeloma; blastoma; seminoma; melanoma; acute myeloid leukaemia.
XX
OS Homo sapiens.
XX
PN WO200259264-A2.
XX
PD 01-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049440.
XX
PR 29-DEC-2000; 2000US-00751181.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Hagai Y, Lazarovits J, Guy R, Lipschitz O, Szanton E, Levanon A;
PI Plaksin D, Peretz T;
XX
DR WPI; 2002-619166/66.
XX
PT Novel peptide/polypeptide for cancer therapy has Fv molecule, construct
PT or fragment, or construct of fragment with enhanced binding
PT characteristics so as to selectively bind target cell in favor of other
PT cells.
XX
PS Claim 4; Page 155-156; 232pp; English.
XX
CC The invention relates to a peptide or polypeptide comprising an Fv
CC molecule, a construct or fragments or a construct of a fragment with
CC enhanced binding characteristics which selectively and/or specifically
CC binds to a target cell in favour of other cells, where binding is
CC primarily determined by a first hypervariable region and Fv is a single
CC chain Fv (scFv) or a disulfide Fv (dsFv). The peptide, optionally in
CC association with or attached, coupled, combined, linked or fused to a
CC pharmaceutical agent, is useful in the manufacture of a medicament, where
CC the medicament has activity against a diseased cell, preferably a cancer
CC cell (selected from carcinoma, sarcoma, leukaemia, adenoma, lymphoma,
CC myeloma, blastoma, seminoma, and melanoma, where the leukaemia cell is an
CC acute myeloid leukaemia cell). The peptide is also useful for preparing a
CC composition for use in inhibiting the growth of a diseased or cancer
CC cell. This sequence represents a human Fv molecule hypervariable region
CC related peptide of the invention
XX
SQ Sequence 277 AA;

Query Match 100.0%; Score 29; DB 5; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 121 MRAPVI 126

RESULT 10
ABG91841
ID ABG91841 standard; protein; 277 AA.
XX
AC ABG91841;
XX
DT 04-DEC-2002 (first entry)
XX
DE Human antibody fragment #25.
XX
KW Human; antibody; epitope; cancer; tumour; cell rolling; inflammation;
KW metastasis; hypervariable region; autoimmune disease; thrombosis;
KW restenosis; leukaemia; inflammatory disease; cardiovascular disease;
KW myocardial infarction; retinopathic disease; abnormal platelet function;
KW sulphated tyrosine-dependent protein-protein interaction.
XX

OS Homo sapiens.
XX
PN WO200253700-A2.
XX
PD 11-JUL-2002.
XX
PF 31-DEC-2001; 2001WO-US049442.
XX
PR 29-DEC-2000; 2000US-00751181.
PR 29-DEC-2000; 2000US-0258948P.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Lazarovits J, Hagai Y, Plaksin D, Vogel T, Nimrod A, Mar-Haim H;
PI Szanthon E, Richter T, Amit B, Kooperman L, Peretz T, Levanon A;
XX
DR WPI; 2002-674776/72.
XX
PT Novel isolated epitope present on cancer cells and important in
PT physiological phenomena such as cell rolling, metastasis and
PT inflammation, for treating autoimmune, inflammatory or cardiovascular
PT diseases, and cancer.
XX
PS Claim 23; Page 233-234; 0pp; English.
XX
CC The invention relates to an isolated epitope present on cancer cells and
CC important in physiological phenomena such as cell rolling, metastasis and
CC inflammation, where the epitope is capable of being bound by an antibody,
CC its antigen-binding fragment or its complex comprising at least one
CC antibody or its binding fragment having a first hypervariable region. The
CC epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
CC disease, thrombosis, restenosis, metastasis, growth and/or replication of
CC tumour or leukaemia cells, increase in number of tumour or leukaemia
CC cells in a patient, cell-cell, cell-matrix, platelet-matrix, platelet-
CC platelet and/or cell-platelet adhesion or aggregation, for increasing
CC mortality of tumour or leukaemia cells, for increasing the susceptibility
CC of diseased cells to damage by anti-disease, anti-cancer or anti-
CC leukaemia agents, or for decreasing the number of tumour or leukaemia
CC cells in a patient, or in the manufacture of a medicament for the above
CC mentioned purposes. The epitopes are useful for diagnosing and treating
CC diseases such as cancer, leukaemia, autoimmune diseases, inflammatory
CC diseases, cardiovascular diseases such as myocardial infarction,
CC retinopathic diseases and other diseases mediated by abnormal platelet
CC function and diseases caused by sulphated tyrosine-dependent protein-
CC protein interactions. This sequence represents a human antibody fragment
CC of the invention
XX
SQ Sequence 277 AA;

Query Match 100.0%; Score 29; DB 5; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 121 MRAPVI 126

RESULT 11
ABG78151
ID ABG78151 standard; protein; 464 AA.
XX
AC ABG78151;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human Fv molecule hypervariable region related peptide #26.
XX
KW Human; Fv molecule; hypervariable region; single chain Fv; cytostatic;
KW disulfide Fv; dsFv; scFv; cancer; carcinoma; sarcoma; leukaemia; adenoma;
KW lymphoma; myeloma; blastoma; seminoma; melanoma; acute myeloid leukaemia.
XX
OS Homo sapiens.
XX
PN WO200259264-A2.
XX
PD 01-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049440.
XX
PR 29-DEC-2000; 2000US-00751181.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Hagai Y, Lazarovits J, Guy R, Lipschitz O, Szanton E, Levanon A;
PI Plaksin D, Peretz T;
XX
DR WPI; 2002-619166/66.
DR N-PSDB; ABS63405.
XX
PT Novel peptide/polypeptide for cancer therapy has Fv molecule, construct
PT or fragment, or construct of fragment with enhanced binding
PT characteristics so as to selectively bind target cell in favor of other
PT cells.
XX
PS Claim 109; Page 93-94; 232pp; English.
XX
CC The invention relates to a peptide or polypeptide comprising an Fv
CC molecule, a construct or fragments or a construct of a fragment with
CC enhanced binding characteristics which selectively and/or specifically
CC binds to a target cell in favour of other cells, where binding is
CC primarily determined by a first hypervariable region and Fv is a single
CC chain Fv (scFv) or a disulfide Fv (dsFv). The peptide, optionally in
CC association with or attached, coupled, combined, linked or fused to a
CC pharmaceutical agent, is useful in the manufacture of a medicament, where
CC the medicament has activity against a diseased cell, preferably a cancer
CC cell (selected from carcinoma, sarcoma, leukaemia, adenoma, lymphoma,
CC myeloma, blastoma, seminoma, and melanoma, where the leukaemia cell is an
CC acute myeloid leukaemia cell). The peptide is also useful for preparing a
CC composition for use in inhibiting the growth of a diseased or cancer
CC cell. This sequence represents a human Fv molecule hypervariable region
CC related peptide of the invention
XX
SQ Sequence 464 AA;

Query Match 100.0%; Score 29; DB 5; Length 464;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 118 MRAPVI 123

RESULT 12
ABG92021
ID ABG92021 standard; protein; 464 AA.
XX
AC ABG92021;
XX
DT 04-DEC-2002 (first entry)
XX
DE Antibody protein #1.
XX
KW Antibody; epitope; cancer; tumour; cell rolling; inflammation;
KW metastasis; hypervariable region; autoimmune disease; thrombosis;
KW restenosis; leukaemia; inflammatory disease; cardiovascular disease;
KW myocardial infarction; retinopathic disease; abnormal platelet function;
KW sulphated tyrosine-dependent protein-protein interaction.
XX
OS Unidentified.
XX
PN WO200253700-A2.
XX
PD 11-JUL-2002.
XX

PF 31-DEC-2001; 2001WO-US049442.
XX
PR 29-DEC-2000; 2000US-00751181.
PR 29-DEC-2000; 2000US-0258948P.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Lazarovits J, Hagai Y, Plaksin D, Vogel T, Nimrod A, Mar-Haim H;
PI Szanthon E, Richter T, Amit B, Kooperman L, Peretz T, Levanon A;
XX
DR WPI; 2002-674776/72.
DR N-PSDB; ABS67759.
XX
PT Novel isolated epitope present on cancer cells and important in
PT physiological phenomena such as cell rolling, metastasis and
PT inflammation, for treating autoimmune, inflammatory or cardiovascular
PT diseases, and cancer.
XX
PS Example 5; Fig 48A; 0pp; English.
XX
CC The invention relates to an isolated epitope present on cancer cells and
CC important in physiological phenomena such as cell rolling, metastasis and
CC inflammation, where the epitope is capable of being bound by an antibody,
CC its antigen-binding fragment or its complex comprising at least one
CC antibody or its binding fragment having a first hypervariable region. The
CC epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
CC disease, thrombosis, restenosis, metastasis, growth and/or replication of
CC tumour or leukaemia cells, increase in number of tumour or leukaemia
CC cells in a patient, cell-cell, cell-matrix, platelet-matrix, platelet-
CC platelet and/or cell-platelet adhesion or aggregation, for increasing
CC mortality of tumour or leukaemia cells, for increasing the susceptibility
CC of diseased cells to damage by anti-disease, anti-cancer or anti-
CC leukaemia agents, or for decreasing the number of tumour or leukaemia
CC cells in a patient, or in the manufacture of a medicament for the above
CC mentioned purposes. The epitopes are useful for diagnosing and treating
CC diseases such as cancer, leukaemia, autoimmune diseases, inflammatory
CC diseases, cardiovascular diseases such as myocardial infarction,
CC retinopathic diseases and other diseases mediated by abnormal platelet
CC function and diseases caused by sulphated tyrosine-dependent protein-
CC protein interactions. This sequence represents an antibody protein of the
CC invention
XX
SQ Sequence 464 AA;

Query Match 100.0%; Score 29; DB 5; Length 464;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 118 MRAPVI 123

RESULT 13
ABG91842
ID ABG91842 standard; protein; 464 AA.
XX
AC ABG91842;
XX
DT 04-DEC-2002 (first entry)
XX
DE Human antibody fragment #26.
XX
KW Human; antibody; epitope; cancer; tumour; cell rolling; inflammation;
KW metastasis; hypervariable region; autoimmune disease; thrombosis;
KW restenosis; leukaemia; inflammatory disease; cardiovascular disease;
KW myocardial infarction; retinopathic disease; abnormal platelet function;
KW sulphated tyrosine-dependent protein-protein interaction.
XX
OS Homo sapiens.
XX
PN WO200253700-A2.
XX

PD 11-JUL-2002.
XX
PF 31-DEC-2001; 2001WO-US049442.
XX
PR 29-DEC-2000; 2000US-00751181.
PR 29-DEC-2000; 2000US-0258948P.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Lazarovits J, Hagai Y, Plaksin D, Vogel T, Nimrod A, Mar-Haim H;
PI Szanthon E, Richter T, Amit B, Kooperman L, Peretz T, Levanon A;
XX
DR WPI; 2002-674776/72.
XX
PT Novel isolated epitope present on cancer cells and important in
PT physiological phenomena such as cell rolling, metastasis and
PT inflammation, for treating autoimmune, inflammatory or cardiovascular
PT diseases, and cancer.
XX
PS Disclosure; Page 234-235; 0pp; English.
XX
CC The invention relates to an isolated epitope present on cancer cells and
CC important in physiological phenomena such as cell rolling, metastasis and
CC inflammation, where the epitope is capable of being bound by an antibody,
CC its antigen-binding fragment or its complex comprising at least one
CC antibody or its binding fragment having a first hypervariable region. The
CC epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
CC disease, thrombosis, restenosis, metastasis, growth and/or replication of
CC tumour or leukaemia cells, increase in number of tumour or leukaemia
CC cells in a patient, cell-cell, cell-matrix, platelet-matrix, platelet-
CC platelet and/or cell-platelet adhesion or aggregation, for increasing
CC mortality of tumour or leukaemia cells, for increasing the susceptibility
CC of diseased cells to damage by anti-disease, anti-cancer or anti-
CC leukaemia agents, or for decreasing the number of tumour or leukaemia
CC cells in a patient, or in the manufacture of a medicament for the above
CC mentioned purposes. The epitopes are useful for diagnosing and treating
CC diseases such as cancer, leukaemia, autoimmune diseases, inflammatory
CC diseases, cardiovascular diseases such as myocardial infarction,
CC retinopathic diseases and other diseases mediated by abnormal platelet
CC function and diseases caused by sulphated tyrosine-dependent protein-
CC protein interactions. This sequence represents a human antibody fragment
CC of the invention
XX
SQ Sequence 464 AA;

Query Match 100.0%; Score 29; DB 5; Length 464;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 118 MRAPVI 123

RESULT 14
AAU42688
ID AAU42688 standard; protein; 133 AA.
XX
AC AAU42688;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #3584.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX

PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59518.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 3883; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 133 AA;

Query Match 96.6%; Score 28; DB 4; Length 133;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||:|
Db 59 MRAPII 64

RESULT 15
ABM39207
ID ABM39207 standard; protein; 133 AA.
XX
AC ABM39207;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #3883.
XX
KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64447.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 3883; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 133 AA;

Query Match 96.6%; Score 28; DB 6; Length 133;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||:|
Db 59 MRAPII 64

Search completed: September 18, 2004, 03:42:51
Job time : 46.4286 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 03:39:30 ; Search time 13.9286 Seconds
(without alignments)
22.239 Million cell updates/sec

Title: US-10-029-926B-8
Perfect score: 29
Sequence: 1 MRAPVI 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	96.6	973	4	US-09-252-991A-23944	Sequence 23944, A
2	26	89.7	286	4	US-09-711-164-385	Sequence 385, App
3	26	89.7	350	4	US-09-198-452A-643	Sequence 643, App
4	26	89.7	477	4	US-09-489-039A-8906	Sequence 8906, Ap
5	25	86.2	119	4	US-09-252-991A-28425	Sequence 28425, A
6	25	86.2	120	4	US-09-107-532A-6688	Sequence 6688, Ap
7	25	86.2	222	4	US-09-489-039A-8244	Sequence 8244, Ap
8	25	86.2	223	4	US-09-134-000C-6010	Sequence 6010, Ap
9	25	86.2	233	4	US-09-489-039A-7366	Sequence 7366, Ap
10	25	86.2	239	4	US-09-328-352-5866	Sequence 5866, Ap
11	25	86.2	359	4	US-09-252-991A-19964	Sequence 19964, A
12	25	86.2	459	4	US-09-252-991A-26783	Sequence 26783, A
13	25	86.2	481	4	US-09-489-039A-10515	Sequence 10515, A
14	25	86.2	657	4	US-09-252-991A-27682	Sequence 27682, A
15	25	86.2	962	4	US-09-328-352-7942	Sequence 7942, Ap
16	24	82.8	73	4	US-09-621-976-6392	Sequence 6392, Ap
17	24	82.8	76	4	US-09-636-215-575	Sequence 575, App
18	24	82.8	76	4	US-09-685-166A-575	Sequence 575, App
19	24	82.8	76	4	US-09-685-166A-888	Sequence 888, App
20	24	82.8	82	4	US-09-134-001C-4676	Sequence 4676, Ap
21	24	82.8	89	4	US-09-134-001C-4563	Sequence 4563, Ap
22	24	82.8	93	4	US-09-938-806A-8	Sequence 8, Appli
23	24	82.8	125	4	US-09-199-637A-403	Sequence 403, App
24	24	82.8	125	4	US-09-269-410-11	Sequence 11, Appl
25	24	82.8	162	4	US-08-801-742-3	Sequence 3, Appli
26	24	82.8	192	1	US-08-208-008C-9	Sequence 9, Appli
27	24	82.8	220	4	US-09-489-039A-8353	Sequence 8353, Ap
28	24	82.8	247	2	US-07-885-089B-2	Sequence 2, Appli
29	24	82.8	247	2	US-07-885-089B-8	Sequence 8, Appli
30	24	82.8	252	2	US-07-885-089B-7	Sequence 7, Appli
31	24	82.8	252	4	US-09-976-594-582	Sequence 582, App
32	24	82.8	253	4	US-09-134-001C-5435	Sequence 5435, Ap
33	24	82.8	269	4	US-09-252-991A-31653	Sequence 31653, A
34	24	82.8	324	4	US-09-489-039A-11252	Sequence 11252, A
35	24	82.8	356	4	US-09-134-001C-3408	Sequence 3408, Ap
36	24	82.8	363	3	US-08-912-560-2	Sequence 2, Appli
37	24	82.8	409	1	US-08-403-545-2	Sequence 2, Appli
38	24	82.8	409	3	US-08-404-381-2	Sequence 2, Appli
39	24	82.8	410	3	US-09-258-754-451	Sequence 451, App
40	24	82.8	411	3	US-09-258-754-448	Sequence 448, App
41	24	82.8	429	4	US-09-252-991A-30995	Sequence 30995, A
42	24	82.8	453	4	US-09-543-681A-5855	Sequence 5855, Ap
43	24	82.8	461	4	US-09-489-039A-8460	Sequence 8460, Ap
44	24	82.8	464	4	US-09-252-991A-16660	Sequence 16660, A
45	24	82.8	478	4	US-09-273-871A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-252-991A-23944
; Sequence 23944, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23944
; LENGTH: 973
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23944

Query Match 96.6%; Score 28; DB 4; Length 973;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRAPVI 6
              |||||:
Db        387 MRAPVV 392
```

RESULT 2
US-09-711-164-385
; Sequence 385, Application US/09711164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERE
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711,164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 385
; LENGTH: 286
; TYPE: PRT

```
;   ORGANISM: Escherichia coli
US-09-711-164-385

  Query Match             89.7%;  Score 26;  DB 4;  Length 286;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||:|||
Db         41 MRSPVI 46


RESULT 3
US-09-198-452A-643
; Sequence 643, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
;  APPLICANT: Griffais, R.
;  TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
;  TITLE OF INVENTION:  thereof and uses thereof, in particular for the diagnosis, preve
;  TITLE OF INVENTION:  and treatment of infection
;  FILE REFERENCE: 9710-003-999
;  CURRENT APPLICATION NUMBER: US/09/198,452A
;  CURRENT FILING DATE:  1998-11-24
;  NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 643
;   LENGTH: 350
;   TYPE: PRT
;   ORGANISM: Chlamydia pneumoniae
US-09-198-452A-643

  Query Match             89.7%;  Score 26;  DB 4;  Length 350;
  Best Local Similarity   83.3%;  Pred. No. 2.7e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              :|||||
Db        109 LRAPVI 114


RESULT 4
US-09-489-039A-8906
; Sequence 8906, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 8906
;   LENGTH: 477
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8906

  Query Match             89.7%;  Score 26;  DB 4;  Length 477;
  Best Local Similarity   83.3%;  Pred. No. 3.8e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |:||||
Db        269 MKAPVI 274


RESULT 5
US-09-252-991A-28425
; Sequence 28425, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 28425
;   LENGTH: 119
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28425

  Query Match             86.2%;  Score 25;  DB 4;  Length 119;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |||| |
Db         63 MRAPTI 68


RESULT 6
US-09-107-532A-6688
; Sequence 6688, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
;             APPLICATION NUMBER: 60/051571
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (781)893-5007
;             TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 6688:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 120 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ORIGINAL SOURCE:
;             ORGANISM: Enterococcus faecium
;        FEATURE:
```

```
;             NAME/KEY:  misc_feature
;             LOCATION:  (B) LOCATION 1...120
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6688:
US-09-107-532A-6688

  Query Match             86.2%;  Score 25;  DB 4;  Length 120;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |:||:|
Db         17 MKAPII 22


RESULT 7
US-09-489-039A-8244
; Sequence 8244, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 8244
;   LENGTH: 222
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8244

  Query Match             86.2%;  Score 25;  DB 4;  Length 222;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPV 5
              |||||
Db         24 MRAPV 28


RESULT 8
US-09-134-000C-6010
; Sequence 6010, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6010
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-6010

  Query Match             86.2%;  Score 25;  DB 4;  Length 223;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPV 5
              |||||
Db        127 MRAPV 131


RESULT 9
US-09-489-039A-7366
; Sequence 7366, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 7366
;   LENGTH: 233
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7366

  Query Match             86.2%;  Score 25;  DB 4;  Length 233;
  Best Local Similarity   100.0%;  Pred. No. 3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPV 5
              |||||
Db        173 MRAPV 177


RESULT 10
US-09-328-352-5866
; Sequence 5866, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5866
;   LENGTH: 239
;   TYPE: PRT
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-5866

  Query Match             86.2%;  Score 25;  DB 4;  Length 239;
  Best Local Similarity   66.7%;  Pred. No. 3.1e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |:|||:
Db        140 MKAPVV 145


RESULT 11
US-09-252-991A-19964
; Sequence 19964, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 19964
```

```
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19964

  Query Match             86.2%;  Score 25;  DB 4;  Length 359;
  Best Local Similarity   66.7%;  Pred. No. 4.8e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              :||||:
Db        212 LRAPVV 217


RESULT 12
US-09-252-991A-26783
; Sequence 26783, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 26783
;   LENGTH: 459
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26783

  Query Match             86.2%;  Score 25;  DB 4;  Length 459;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPV 5
              |||||
Db         38 MRAPV 42


RESULT 13
US-09-489-039A-10515
; Sequence 10515, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 10515
;   LENGTH: 481
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10515

  Query Match             86.2%;  Score 25;  DB 4;  Length 481;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPV 5
              |||||
Db        131 MRAPV 135


RESULT 14
US-09-252-991A-27682
; Sequence 27682, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 27682
;   LENGTH: 657
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27682

  Query Match             86.2%;  Score 25;  DB 4;  Length 657;
  Best Local Similarity   66.7%;  Pred. No. 9.1e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              :||||:
Db        332 LRAPVV 337


RESULT 15
US-09-328-352-7942
; Sequence 7942, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7942
;   LENGTH: 962
;   TYPE: PRT
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-7942

  Query Match             86.2%;  Score 25;  DB 4;  Length 962;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+03;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |||| |
Db        471 MRAPAI 476


Search completed: September 18, 2004, 03:48:46
Job time : 15.9286 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 03:46:39 ; Search time 238.286 Seconds
(without alignments)
8.086 Million cell updates/sec

Title: US-10-029-926B-8
Perfect score: 29
Sequence: 1 MRAPVI 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	6	12	US-10-029-926B-8	Sequence 8, Appli
2	29	100.0	6	15	US-10-032-037B-8	Sequence 8, Appli
3	29	100.0	6	15	US-10-029-988B-8	Sequence 8, Appli
4	29	100.0	6	15	US-10-032-423A-8	Sequence 8, Appli
5	29	100.0	266	15	US-10-032-037B-204	Sequence 204, App
6	29	100.0	266	15	US-10-029-988B-204	Sequence 204, App
7	29	100.0	266	15	US-10-032-423A-204	Sequence 204, App
8	29	100.0	277	12	US-10-029-926B-25	Sequence 25, Appl
9	29	100.0	277	15	US-10-032-037B-25	Sequence 25, Appl
10	29	100.0	277	15	US-10-029-988B-25	Sequence 25, Appl
11	29	100.0	277	15	US-10-032-423A-25	Sequence 25, Appl
12	29	100.0	442	15	US-10-369-493-7871	Sequence 7871, Ap
13	29	100.0	464	12	US-10-029-926B-26	Sequence 26, Appl
14	29	100.0	464	15	US-10-032-037B-26	Sequence 26, Appl
15	29	100.0	464	15	US-10-029-988B-26	Sequence 26, Appl
16	29	100.0	464	15	US-10-032-423A-26	Sequence 26, Appl
17	28	96.6	86	12	US-10-424-599-253048	Sequence 253048,
18	28	96.6	123	12	US-10-424-599-150638	Sequence 150638,
19	28	96.6	438	9	US-09-908-419-2	Sequence 2, Appli
20	28	96.6	438	10	US-09-907-187-2	Sequence 2, Appli
21	28	96.6	438	14	US-10-056-790-2	Sequence 2, Appli
22	28	96.6	438	14	US-10-056-790-36	Sequence 36, Appl
23	28	96.6	470	14	US-10-056-790-46	Sequence 46, Appl
24	28	96.6	541	9	US-09-738-626-3996	Sequence 3996, Ap
25	27	93.1	99	12	US-10-424-599-272558	Sequence 272558,
26	26	89.7	112	16	US-10-767-701-42410	Sequence 42410, A
27	26	89.7	226	14	US-10-032-585-7418	Sequence 7418, Ap
28	26	89.7	286	12	US-10-282-122A-43216	Sequence 43216, A
29	26	89.7	286	14	US-10-287-274-385	Sequence 385, App
30	26	89.7	286	15	US-10-369-493-23563	Sequence 23563, A
31	26	89.7	289	12	US-10-425-114-69148	Sequence 69148, A
32	26	89.7	290	12	US-10-282-122A-56193	Sequence 56193, A
33	26	89.7	319	12	US-10-425-114-55531	Sequence 55531, A
34	26	89.7	345	16	US-10-767-701-45307	Sequence 45307, A
35	26	89.7	350	15	US-10-289-762-643	Sequence 643, App
36	26	89.7	405	14	US-10-017-161-2090	Sequence 2090, Ap
37	26	89.7	405	15	US-10-292-798-1736	Sequence 1736, Ap
38	26	89.7	514	12	US-10-425-114-53835	Sequence 53835, A
39	26	89.7	530	12	US-10-425-114-60637	Sequence 60637, A
40	26	89.7	1428	16	US-10-437-963-144989	Sequence 144989,
41	25	86.2	58	16	US-10-437-963-137085	Sequence 137085,
42	25	86.2	63	16	US-10-767-701-48548	Sequence 48548, A
43	25	86.2	75	12	US-10-424-599-196866	Sequence 196866,
44	25	86.2	104	16	US-10-437-963-164399	Sequence 164399,
45	25	86.2	110	16	US-10-437-963-158848	Sequence 158848,

ALIGNMENTS

RESULT 1
US-10-029-926B-8
; Sequence 8, Application US/10029926B
; Publication No. US20040073011A1
; GENERAL INFORMATION:
; APPLICANT: HAGAY, et al.
; TITLE OF INVENTION: SPECIFIC HUMAN ANTIBODIES FOR SELECTIVE CANCER THERAPY
; FILE REFERENCE: 10793/50
; CURRENT APPLICATION NUMBER: US/10/029,926B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-926B-8

Query Match 100.0%; Score 29; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 1 MRAPVI 6

RESULT 2
US-10-032-037B-8
; Sequence 8, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/44

```
;  CURRENT APPLICATION NUMBER: US/10/032,037B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-032-037B-8

  Query Match            100.0%;  Score 29;  DB 15;  Length 6;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+06;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db          1 MRAPVI 6


RESULT 3
US-10-029-988B-8
; Sequence 8, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/46
;  CURRENT APPLICATION NUMBER: US/10/029,988B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-029-988B-8

  Query Match            100.0%;  Score 29;  DB 15;  Length 6;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+06;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db          1 MRAPVI 6


RESULT 4
US-10-032-423A-8
; Sequence 8, Application US/10032423A
; Publication No. US20040002450A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/45
;  CURRENT APPLICATION NUMBER: US/10/032,423A
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 12/29/2000
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 6
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-032-423A-8

  Query Match            100.0%;  Score 29;  DB 15;  Length 6;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+06;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db          1 MRAPVI 6


RESULT 5
US-10-032-037B-204
; Sequence 204, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/44
;  CURRENT APPLICATION NUMBER: US/10/032,037B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 204
;   LENGTH: 266
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-032-037B-204

  Query Match            100.0%;  Score 29;  DB 15;  Length 266;
  Best Local Similarity  100.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db        121 MRAPVI 126


RESULT 6
US-10-029-988B-204
; Sequence 204, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/46
;  CURRENT APPLICATION NUMBER: US/10/029,988B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 204
;   LENGTH: 266
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-029-988B-204

  Query Match            100.0%;  Score 29;  DB 15;  Length 266;
  Best Local Similarity  100.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db        121 MRAPVI 126


RESULT 7
US-10-032-423A-204
; Sequence 204, Application US/10032423A
; Publication No. US20040002450A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/45
;  CURRENT APPLICATION NUMBER: US/10/032,423A
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 12/29/2000
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 204
;   LENGTH: 266
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-032-423A-204

  Query Match             100.0%;  Score 29;  DB 15;  Length 266;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db        121 MRAPVI 126


RESULT 8
US-10-029-926B-25
; Sequence 25, Application US/10029926B
; Publication No. US20040073011A1
; GENERAL INFORMATION:
;  APPLICANT: HAGAY, et al.
;  TITLE OF INVENTION: SPECIFIC HUMAN ANTIBODIES FOR SELECTIVE CANCER THERAPY
;  FILE REFERENCE: 10793/50
;  CURRENT APPLICATION NUMBER: US/10/029,926B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 12/29/2000
;  NUMBER OF SEQ ID NOS: 203
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
;   LENGTH: 277
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-029-926B-25

  Query Match             100.0%;  Score 29;  DB 12;  Length 277;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db        121 MRAPVI 126


RESULT 9
US-10-032-037B-25
; Sequence 25, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/44
;  CURRENT APPLICATION NUMBER: US/10/032,037B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
;   LENGTH: 277
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-032-037B-25

  Query Match             100.0%;  Score 29;  DB 15;  Length 277;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db        121 MRAPVI 126


RESULT 10
US-10-029-988B-25
; Sequence 25, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/46
;  CURRENT APPLICATION NUMBER: US/10/029,988B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
;   LENGTH: 277
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-029-988B-25

  Query Match             100.0%;  Score 29;  DB 15;  Length 277;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db        121 MRAPVI 126


RESULT 11
US-10-032-423A-25
; Sequence 25, Application US/10032423A
; Publication No. US20040002450A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/45
;  CURRENT APPLICATION NUMBER: US/10/032,423A
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 12/29/2000
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
;   LENGTH: 277
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-032-423A-25

  Query Match             100.0%;  Score 29;  DB 15;  Length 277;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db        121 MRAPVI 126


RESULT 12
US-10-369-493-7871
```

; Sequence 7871, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7871
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Rhodobacter sphaeroides
US-10-369-493-7871

Query Match 100.0%; Score 29; DB 15; Length 442;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 77 MRAPVI 82

RESULT 13
US-10-029-926B-26
; Sequence 26, Application US/10029926B
; Publication No. US20040073011A1
; GENERAL INFORMATION:
; APPLICANT: HAGAY, et al.
; TITLE OF INVENTION: SPECIFIC HUMAN ANTIBODIES FOR SELECTIVE CANCER THERAPY
; FILE REFERENCE: 10793/50
; CURRENT APPLICATION NUMBER: US/10/029,926B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-926B-26

Query Match 100.0%; Score 29; DB 12; Length 464;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 118 MRAPVI 123

RESULT 14
US-10-032-037B-26
; Sequence 26, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/44
; CURRENT APPLICATION NUMBER: US/10/032,037B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-037B-26

Query Match 100.0%; Score 29; DB 15; Length 464;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 118 MRAPVI 123

RESULT 15
US-10-029-988B-26
; Sequence 26, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/46
; CURRENT APPLICATION NUMBER: US/10/029,988B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-988B-26

Query Match 100.0%; Score 29; DB 15; Length 464;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||||||
Db 118 MRAPVI 123

Search completed: September 18, 2004, 04:20:41
Job time : 239.286 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 03:33:04 ; Search time 12.2143 Seconds
(without alignments)
47.252 Million cell updates/sec

Title: US-10-029-926B-8
Perfect score: 29
Sequence: 1 MRAPVI 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	249	2	H87675	hydrolase, alpha/b
2	29	100.0	442	2	JC4733	probable H+-transp
3	29	100.0	629	2	T37255	acetylcholinestera
4	28	96.6	425	2	S33045	hypothetical prote
5	27	93.1	247	2	H69030	coenzyme PQQ synth
6	26	89.7	286	2	E65103	tagatose-bisphosph
7	26	89.7	286	2	A91131	tagatose-1,6-bisph
8	26	89.7	286	2	A85976	tagatose-bisphosph
9	26	89.7	327	2	E72057	ferrochelatase CP0
10	26	89.7	327	2	H86565	ferrochetalase [im
11	26	89.7	450	1	S13730	pmbA protein - Esc
12	26	89.7	450	2	D91280	maturation of anti
13	26	89.7	450	2	D86121	maturation of anti
14	26	89.7	450	2	AE1056	probable PmbA prot
15	25	86.2	138	2	T24916	hypothetical prote
16	25	86.2	147	2	C95907	probable protein i
17	25	86.2	147	2	F95924	hypothetical prote
18	25	86.2	149	2	C72419	conserved hypothet
19	25	86.2	154	2	A82812	conserved hypothet
20	25	86.2	163	2	A87412	hypothetical prote
21	25	86.2	195	2	F75399	antibiotic resista
22	25	86.2	196	2	G75405	probable amidotran
23	25	86.2	203	2	C84409	imidazoleglycerol-
24	25	86.2	222	2	S56195	probable membrane
25	25	86.2	222	2	S63320	probable membrane
26	25	86.2	224	2	S55081	hypothetical prote
27	25	86.2	227	2	C53304	transfer protein C
28	25	86.2	240	2	E69004	hypothetical prote
29	25	86.2	252	2	A88508	protein H14A12.4 [
30	25	86.2	261	2	AC3507	hisF protein [impo
31	25	86.2	266	2	AC2407	tryptophan synthas
32	25	86.2	284	2	AD1712	fructose-1,6-bipho
33	25	86.2	284	2	AF1341	fructose-1,6-bipho
34	25	86.2	328	2	G96024	conserved hypothet
35	25	86.2	332	2	C81285	glyceraldehyde 3-p
36	25	86.2	359	2	AI0112	fructose-bisphosph
37	25	86.2	362	2	F90200	hypothetical prote
38	25	86.2	398	2	B70752	hypothetical prote
39	25	86.2	422	2	AE3394	lipoprotein releas
40	25	86.2	422	2	E83083	probable two-compo
41	25	86.2	430	2	T16715	hypothetical prote
42	25	86.2	446	2	A84940	pmbA protein [impo
43	25	86.2	463	2	AC0977	L-seryl-tRNASec se
44	25	86.2	487	2	T27353	hypothetical prote
45	25	86.2	489	2	JC4787	shaw protein - Cal

ALIGNMENTS

RESULT 1
H87675
hydrolase, alpha/beta hydrolase fold family [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: H87675
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: H87675
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <STO>
A;Cross-references: GB:AE005673; NID:g13425158; PIDN:AAK25404.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC3442

Query Match 100.0%; Score 29; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRAPVI 6
              ||||||
Db          1 MRAPVI 6
```

RESULT 2
JC4733
probable H+-transporting two-sector ATPase (EC 3.6.3.14), flagellum-specific - Rhodobac
C;Species: Rhodobacter sphaeroides
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 03-Jun-2002
C;Accession: JC4733
R;Ballado, T.; Campos, A.; Camarena, L.; Dreyfus, G.
Gene 170, 69-72, 1996
A;Title: Flagellar genes from Rhodobacter sphaeroides are homologous to genes of the fl
A;Reference number: JC4733; MUID:96200857; PMID:8621091
A;Accession: JC4733
A;Molecule type: DNA
A;Residues: 1-442 <BAL>
A;Cross-references: GB:U31090; NID:g1518877; PIDN:AAB07344.1; PID:g1293928
C;Genetics:
A;Gene: fliI
C;Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase a
C;Keywords: ATP; flagellum; hydrolase; nucleotide binding; P-loop
F;168-175/Region: nucleotide-binding motif A (P-loop)
F;191-362/Domain: H+-transporting ATP synthase alpha chain homology <ATP>
F;191-207/Region: ATP-binding #status predicted
F;242-258/Domain: beta chain #status predicted <BET>

Query Match 100.0%; Score 29; DB 2; Length 442;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRAPVI 6
              ||||||
Db         77 MRAPVI 82
```

RESULT 3
T37255
acetylcholinesterase (EC 3.1.1.7) 2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C;Accession: T37255
R;Grauso, M.; Culetto, E.; Combes, D.; Fedon, Y.; Toutant, J.P.; Arpagaus, M.
FEBS Lett. 424, 279-284, 1998
A;Title: Existence of four acetylcholinesterase genes in the nematodes Caenorhabditis el
A;Reference number: Z21648; MUID:98198570; PMID:9539167
A;Accession: T37255
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-629 <GRA>
A;Cross-references: EMBL:AF025378; NID:g5148937; PIDN:AAC14016.2; PID:g5091489
A;Experimental source: strain N2
C;Genetics:
A;Gene: ace-2
A;Map position: I
C;Superfamily: cholinesterase; cholinesterase homology
C;Keywords: carboxylic ester hydrolase

Query Match 100.0%; Score 29; DB 2; Length 629;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRAPVI 6
              ||||||
Db          1 MRAPVI 6
```

RESULT 4
S33045
hypothetical protein - human herpesvirus 4
C;Species: human herpesvirus 4, Epstein-Barr virus
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Jun-2000
C;Accession: S33045
R;Farrell, P.J.
submitted to the EMBL Data Library, March 1988
A;Reference number: S32973
A;Accession: S33045
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-425 <FAR>
A;Cross-references: EMBL:V01555; NID:g59074; PIDN:CAA24796.1; PID:g1334904
C;Superfamily: equine herpesvirus 2 hypothetical protein 23

Query Match 96.6%; Score 28; DB 2; Length 425;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRAPVI 6
              ||||:|
Db         87 MRAPII 92
```

RESULT 5
H69030
coenzyme PQQ synthesis protein III - Methanobacterium thermoautotrophicum (strain Delta
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: H69030
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: H69030
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-247 <MTH>
A;Cross-references: GB:AE000890; GB:AE000666; NID:g2622331; PIDN:AAB85716.1; PID:g262233
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1227
A;Start codon: GTG
C;Keywords: iron; metalloprotein
F;34,38,41/Binding site: iron (Cys) #status predicted

Query Match 93.1%; Score 27; DB 2; Length 247;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRAPVI 6
              ||||::
Db          4 MRAPIV 9
```

RESULT 6
E65103
tagatose-bisphosphate aldolase agaY (EC 4.1.2.-) - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: E65103
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: E65103
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-286 <BLAT>
A;Cross-references: GB:AE000395; GB:U00096; NID:g1789524; PIDN:AAC76171.1; PID:g1789526;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: agaY
C;Superfamily: fructose-bisphosphate aldolase II
C;Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 89.7%; Score 26; DB 2; Length 286;
Best Local Similarity 83.3%; Pred. No. 87;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRAPVI 6
              ||:|||
Db         41 MRSPVI 46
```

RESULT 7
A91131
tagatose-1,6-bisphosphate aldolase [imported] - Escherichia coli (strain O157:H7, substr
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: A91131
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: A91131
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-286 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB37440.1; PID:g13363490; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:
A;Gene: ECs4017
C;Superfamily: fructose-bisphosphate aldolase II

Query Match 89.7%; Score 26; DB 2; Length 286;
Best Local Similarity 83.3%; Pred. No. 87;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||:|||
Db 41 MRSPVI 46

RESULT 8
A85976
tagatose-bisphosphate aldolase 2 [imported] - Escherichia coli (strain O157:H7, substrai
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: A85976
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: A85976
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-286 <STO>
A;Cross-references: GB:AE005174; NID:g12517735; PIDN:AAG58269.1; GSPDB:GN00145; UWGP:Z44
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: agaY
C;Superfamily: fructose-bisphosphate aldolase II

Query Match 89.7%; Score 26; DB 2; Length 286;
Best Local Similarity 83.3%; Pred. No. 87;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
||:|||
Db 41 MRSPVI 46

RESULT 9
E72057
ferrochelatase CP0144 [imported] - Chlamydophila pneumoniae (strains CWL029 and AR39)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C;Accession: E72057; F81609
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388
A;Accession: E72057
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-327 <ARN>
A;Cross-references: GB:AE001645; GB:AE001363; NID:g4376896; PIDN:AAD18742.1; PID:g437689
A;Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: F81609
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-327 <REA>
A;Cross-references: GB:AE002175; GB:AE002161; NID:g7189069; PIDN:AAF38026.1; PID:g718907
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: hemZ; CP0144
C;Superfamily: ferrochelatase

Query Match 89.7%; Score 26; DB 2; Length 327;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
:|||||
Db 86 LRAPVI 91

RESULT 10
H86565
ferrochetalase [imported] - Chlamydophila pneumoniae (strain J138)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C;Accession: H86565
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349; PMID:10871362
A;Accession: H86565
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-327 <STO>
A;Cross-references: GB:BA000008; NID:g8978975; PIDN:BAA98810.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: hemZ

Query Match 89.7%; Score 26; DB 2; Length 327;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
:|||||
Db 86 LRAPVI 91

RESULT 11
S13730
pmbA protein - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 29-Jan-1999 #sequence_revision 29-Jan-1999 #text_change 01-Mar-2002
C;Accession: S13730; S56461; F65235
R;Rodriguez-Sainz, M.C.; Hernandez-Chico, C.; Moreno, F.
Mol. Microbiol. 4, 1921-1932, 1990
A;Title: Molecular characterization of pmbA, an Escherichia coli chromosomal gene requi
A;Reference number: S13730; MUID:91186828; PMID:2082149
A;Accession: S13730
A;Molecule type: DNA
A;Residues: 1-450 <ROD>
A;Cross-references: EMBL:X54152; NID:g42439; PIDN:CAA38091.1; PID:g42440
R;Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 23, 2105-2119, 1995
A;Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92
A;Reference number: S56314; MUID:95334362; PMID:7610040
A;Accession: S56461
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-450 <BUR>
A;Cross-references: EMBL:U14003; NID:g1263172; PIDN:AAA97132.1; PID:g537077
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: F65235
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-450 <BLAT>
A;Cross-references: GB:AE000494; GB:U00096; NID:g1790670; PIDN:AAC77192.1; PID:g1790682
A;Experimental source: strain K-12, substrain MG1655

C;Genetics:
A;Gene: pmbA
A;Map position: 96 min
C;Superfamily: Escherichia coli pmbA protein

  Query Match 89.7%; Score 26; DB 1; Length 450;
  Best Local Similarity 83.3%; Pred. No. 1.4e+02;
  Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
|:||||
Db 242 MKAPVI 247

RESULT 12
D91280
maturation of antibiotic MccB17 [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: D91280
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D91280
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-450 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB38635.1; PID:g13364689; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs5212
C;Superfamily: Escherichia coli pmbA protein

  Query Match 89.7%; Score 26; DB 2; Length 450;
  Best Local Similarity 83.3%; Pred. No. 1.4e+02;
  Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
|:||||
Db 242 MKAPVI 247

RESULT 13
D86121
maturation of antibiotic MccB17, see tld genes [imported] - Escherichia coli (strain O15
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: D86121
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D86121
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-450 <STO>
A;Cross-references: GB:AE005174; NID:g12519236; PIDN:AAG59432.1; GSPDB:GN00145; UWGP:Z58
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: pmbA
C;Superfamily: Escherichia coli pmbA protein

  Query Match 89.7%; Score 26; DB 2; Length 450;
  Best Local Similarity 83.3%; Pred. No. 1.4e+02;
  Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
|:||||
Db 242 MKAPVI 247

RESULT 14
AE1056
probable PmbA protein pmbA [imported] - Salmonella enterica subsp. enterica serovar Typh
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AE1056
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AE1056
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-450 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD06898.1; PID:g16505546; GSPDB:GN00176
C;Genetics:
A;Gene: pmbA
C;Superfamily: Escherichia coli pmbA protein

  Query Match 89.7%; Score 26; DB 2; Length 450;
  Best Local Similarity 83.3%; Pred. No. 1.4e+02;
  Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6
|:||||
Db 242 MKAPVI 247

RESULT 15
T24916
hypothetical protein T14G10.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T24916
R;Wild, A.
submitted to the EMBL Data Library, January 1996
A;Reference number: Z19954
A;Accession: T24916
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-138 <WIL>
A;Cross-references: EMBL:Z68880; PIDN:CAA93096.1; GSPDB:GN00022; CESP:T14G10.4
A;Experimental source: clone T14G10
C;Genetics:
A;Gene: CESP:T14G10.4
A;Map position: 4
A;Introns: 52/3; 108/2

  Query Match 86.2%; Score 25; DB 2; Length 138;
  Best Local Similarity 100.0%; Pred. No. 73;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPV 5
|||||
Db 1 MRAPV 5

Search completed: September 18, 2004, 03:47:30
Job time : 15.2143 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 02:29:47 ; Search time 8.78571 Seconds
(without alignments)
35.560 Million cell updates/sec

Title: US-10-029-926B-8
Perfect score: 29
Sequence: 1 MRAPVI 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	96.6	164	1	RHB1_RAT	O88779 rattus norv
2	28	96.6	373	1	RHB1_MOUSE	Q8vc82 mus musculu
3	28	96.6	425	1	YTR1_EBV	P30119 epstein-bar
4	28	96.6	438	1	RHB1_HUMAN	O75783 homo sapien
5	27	93.1	535	1	TLE1_BRARE	O13168 brachydanio
6	26	89.7	250	1	PDXJ_BRAJA	O69158 bradyrhizob
7	26	89.7	286	1	AGAY_ECOLI	P42908 escherichia
8	26	89.7	327	1	HEMZ_CHLPN	Q9z7v1 chlamydia p
9	26	89.7	450	1	PMBA_ECOLI	P24231 escherichia
10	25	86.2	95	1	CH10_XANAC	Q8rit8 xanthomonas
11	25	86.2	215	1	HIS5_STRAW	Q82aa2 streptomyce
12	25	86.2	222	1	YFG0_YEAST	P43544 saccharomyc
13	25	86.2	222	1	YN74_YEAST	P53823 saccharomyc
14	25	86.2	224	1	YMY5_YEAST	Q03144 saccharomyc
15	25	86.2	258	1	HIS6_RHIME	Q92tb3 rhizobium m
16	25	86.2	261	1	HIS6_BRUME	Q8ye37 brucella me
17	25	86.2	261	1	HIS6_BRUSU	Q8fy07 brucella su
18	25	86.2	398	1	Y106_MYCTU	Q10899 mycobacteri
19	25	86.2	446	1	PMBA_BUCAI	P57191 buchnera ap
20	25	86.2	463	1	SELA_SALTI	Q8z2d8 salmonella
21	25	86.2	463	1	SELA_SALTY	Q8zl69 salmonella
22	25	86.2	486	1	COBB_STRCO	Q9rj16 streptomyce
23	25	86.2	510	1	INO1_HORVU	O65195 hordeum vul
24	25	86.2	545	1	SYFB_METMA	Q8pta5 methanosarc
25	25	86.2	550	1	VGLE_HSV11	P04488 herpes simp
26	25	86.2	1220	1	IF2P_HUMAN	O60841 homo sapien
27	24	82.8	67	1	Y7K8_BPP22	P14108 bacteriopha
28	24	82.8	109	1	YS88_CAEEL	Q09384 caenorhabdi
29	24	82.8	122	1	INL3_MOUSE	O09107 mus musculu
30	24	82.8	162	1	HRS3_HUMAN	P53816 homo sapien
31	24	82.8	186	1	YG61_METJA	Q59055 methanococc
32	24	82.8	216	1	ALKB_ECOLI	P05050 escherichia
33	24	82.8	216	1	ALKB_SALTY	P37462 salmonella
34	24	82.8	243	1	TPIS_PROMM	Q7v7d2 prochloroco
35	24	82.8	243	1	TPIS_SYNPX	Q7u804 synechococc
36	24	82.8	248	1	TPIS_CLOAB	O52633 clostridium
37	24	82.8	248	1	TPIS_CLOPE	Q8xku1 clostridium
38	24	82.8	252	1	AMPR_HUMAN	P15514 homo sapien
39	24	82.8	252	1	TPIS_LACPL	Q88yh4 lactobacill
40	24	82.8	253	1	TPIS_STAAM	Q9z5c3 staphylococ
41	24	82.8	253	1	TPIS_STAEP	Q8ctd5 staphylococ
42	24	82.8	259	1	IF2C_GALSU	Q08810 galdieria s
43	24	82.8	263	1	HIS6_RHILO	Q98ct1 rhizobium l
44	24	82.8	287	1	NAPH_ECOLI	P33934 escherichia
45	24	82.8	323	1	SOPB_ECOLI	P08867 escherichia

ALIGNMENTS

RESULT 1
RHB1_RAT
ID RHB1_RAT STANDARD; PRT; 164 AA.
AC O88779;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rhomboid-related protein 1 (EC 3.4.21.-) (RRP) (Rhomboid-like protein
DE 1) (Fragment).
GN RHBDL1 OR RHBDL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestinal epithelium;
RX MEDLINE=98324821; PubMed=9662444;
RA Pascall J.C., Brown K.D.;
RT "Characterization of a mammalian cDNA encoding a protein with high
RT sequence similarity to the Drosophila regulatory protein Rhomboid.";
RL FEBS Lett. 429:337-340(1998).
CC -!- FUNCTION: May be involved in regulated intramembrane proteolysis
CC and the subsequent release of functional polypeptides from their
CC membrane anchors (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to peptidase family S54.
CC --
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC --
DR EMBL; Y17258; CAA76716.1; -.
DR MEROPS; S54.950; -.
DR InterPro; IPR002610; Peptidase_S54.
DR Pfam; PF01694; Rhomboid; 1.
KW Hydrolase; Protease; Serine protease; Transmembrane.
FT NON_TER 1 1
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 32 52 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT ACT_SITE 60 60 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 125 125 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT NON_TER 164 164
SQ SEQUENCE 164 AA; 17662 MW; CF62ACE3B6C99210 CRC64;

Query Match 96.6%; Score 28; DB 1; Length 164;
Best Local Similarity 83.3%; Pred. No. 6.3;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAPVI 6

```
              |||||:
Db         52 MRAPVV 57


RESULT 2
RHB1_MOUSE
ID   RHB1_MOUSE     STANDARD;      PRT;   373 AA.
AC   Q8VC82;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Rhomboid-related protein 1 (EC 3.4.21.-) (RRP) (Rhomboid-like protein
DE   1).
GN   RHBDL1 OR RHBDL.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Retina;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May be involved in regulated intramembrane proteolysis
CC       and the subsequent release of functional polypeptides from their
CC       membrane anchors (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -!- SIMILARITY: Belongs to peptidase family S54.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; BC021549; AAH21549.1; -.
DR   MEROPS; S54.950; -.
DR   MGD; MGI:2384891; Rhbdl.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR002610; Peptidase_S54.
DR   Pfam; PF01694; Rhomboid; 1.
KW   Hydrolase; Protease; Serine protease; Transmembrane.
FT   TRANSMEM    131    151       POTENTIAL.
FT   TRANSMEM    196    216       POTENTIAL.
FT   TRANSMEM    219    239       POTENTIAL.
FT   TRANSMEM    243    263       POTENTIAL.
FT   TRANSMEM    275    294       POTENTIAL.
FT   TRANSMEM    307    327       POTENTIAL.
FT   TRANSMEM    340    360       POTENTIAL.
FT   ACT_SITE    199    199       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    247    247       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    312    312       CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ   SEQUENCE   373 AA;  41786 MW;  1F4E538B3A363D2A CRC64;

  Query Match             96.6%;  Score 28;  DB 1;  Length 373;
  Best Local Similarity   83.3%;  Pred. No. 15;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |||||:
Db        239 MRAPVV 244


RESULT 3
YTR1_EBV
ID   YTR1_EBV       STANDARD;      PRT;   425 AA.
AC   P30119;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-APR-1993 (Rel. 25, Last annotation update)
DE   Hypothetical BTRF1 protein.
GN   BTRF1.
OS   Epstein-barr virus (strain B95-8) (Human herpesvirus 4).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Lymphocryptovirus.
OX   NCBI_TaxID=10377;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84270667; PubMed=6087149;
RA   Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA   Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA   Tuffnell P.S., Barrell B.G.;
RT   "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL   Nature 310:207-211(1984).
CC   -!- SIMILARITY: TO HVS-1 GENE 23.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V01555; CAA24796.1; -.
DR   PIR; S33045; S33045.
DR   InterPro; IPR006772; Herpes_BTRF1.
DR   Pfam; PF04682; Herpes_BTRF1; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   425 AA;  46711 MW;  0ECCBE5FD30495BD CRC64;

  Query Match             96.6%;  Score 28;  DB 1;  Length 425;
  Best Local Similarity   83.3%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||:|
Db         87 MRAPII 92


RESULT 4
RHB1_HUMAN
ID   RHB1_HUMAN     STANDARD;      PRT;   438 AA.
AC   O75783; Q9NQ85;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Rhomboid-related protein 1 (EC 3.4.21.-) (RRP) (Rhomboid-like protein
DE   1).
GN   RHBDL1 OR RHBDL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
```

```
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Leukemia;
RX   MEDLINE=98324821; PubMed=9662444;
RA   Pascall J.C., Brown K.D.;
RT   "Characterization of a mammalian cDNA encoding a protein with high
RT   sequence similarity to the Drosophila regulatory protein Rhomboid.";
RL   FEBS Lett. 429:337-340(1998).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=21096910; PubMed=11157797;
RA   Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA   Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA   Higgs D.R.;
RT   "Sequence, structure and pathology of the fully annotated terminal 2
RT   Mb of the short arm of human chromosome 16.";
RL   Hum. Mol. Genet. 10:339-352(2001).
CC   -!- FUNCTION: May be involved in regulated intramembrane proteolysis
CC       and the subsequent release of functional polypeptides from their
CC       membrane anchors (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=O75783-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=O75783-2; Sequence=VSP_005372;
CC   -!- TISSUE SPECIFICITY: Detected in heart, brain, skeletal muscle and
CC       kidney.
CC   -!- SIMILARITY: Belongs to peptidase family S54.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y17108; CAA76629.1; -.
DR   EMBL; AJ272344; CAC00640.1; -.
DR   EMBL; AE006464; AAK61241.1; -.
DR   Genew; HGNC:10007; RHBDL1.
DR   MIM; 603264; -.
DR   MEROPS; S54.950; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005624; C:membrane fraction; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR002610; Peptidase_S54.
DR   Pfam; PF01694; Rhomboid; 1.
KW   Hydrolase; Protease; Serine protease; Transmembrane;
KW   Alternative splicing.
FT   TRANSMEM    196    216       POTENTIAL.
FT   TRANSMEM    262    282       POTENTIAL.
FT   TRANSMEM    284    304       POTENTIAL.
FT   TRANSMEM    308    328       POTENTIAL.
FT   TRANSMEM    340    359       POTENTIAL.
FT   TRANSMEM    372    392       POTENTIAL.
FT   TRANSMEM    405    425       POTENTIAL.
FT   ACT_SITE    264    264       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    312    312       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    377    377       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   VARSPLIC      1     77       MGRVEDGGTTEELEDWDPGTSALPAPGIKQGPREQTGTGPL
FT                                SQKCWEPEPDAPSQPGPALWSRGRARTQALAGGSSL -> M
FT                                DRSSLLQLIQE (in isoform 2).
FT                                /FTId=VSP_005372.
SQ   SEQUENCE   438 AA;  48314 MW;  A7644AD39644A2F6 CRC64;

  Query Match             96.6%;  Score 28;  DB 1;  Length 438;
  Best Local Similarity   83.3%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |||||:
Db        304 MRAPVV 309


RESULT 5
TLE1_BRARE
ID   TLE1_BRARE     STANDARD;      PRT;   535 AA.
AC   O13168;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Groucho 1 protein (Fragment).
GN   GRO1 OR GROUCHO1.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wuelbeck C., Campos-Ortega J.A.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: NUCLEAR EFFECTOR MOLECULE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- SIMILARITY: Contains 6 WD repeats.
CC   -!- SIMILARITY: Belongs to the WD-repeat Groucho/TLE family.
CC       PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y12467; CAA73070.1; -.
DR   EMBL; U96451; AAB57808.1; -.
DR   ZFIN; ZDB-GENE-990415-85; gro1.
DR   InterPro; IPR001680; WD40.
DR   Pfam; PF00400; WD40; 6.
DR   ProDom; PD000018; WD40; 1.
DR   SMART; SM00320; WD40; 7.
DR   PROSITE; PS00678; WD_REPEATS_1; 2.
DR   PROSITE; PS50082; WD_REPEATS_2; 2.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Nuclear protein; Repeat; WD repeat.
FT   NON_TER       1      1
FT   DOMAIN       <1     42       CCN DOMAIN.
FT   DOMAIN        43    214       SER/PRO-RICH.
FT   REPEAT       235    266       WD 1.
FT   REPEAT       293    323       WD 2.
FT   REPEAT       337    367       WD 3.
FT   REPEAT       379    409       WD 4.
FT   REPEAT       461    491       WD 5.
FT   REPEAT       502    532       WD 6.
SQ   SEQUENCE   535 AA;  57360 MW;  A201A32071101363 CRC64;

  Query Match             93.1%;  Score 27;  DB 1;  Length 535;
  Best Local Similarity   83.3%;  Pred. No. 41;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |||||:
Db        192 MRAPVL 197


RESULT 6
PDXJ_BRAJA
ID   PDXJ_BRAJA     STANDARD;      PRT;   250 AA.
AC   O69158;
DT   30-MAY-2000 (Rel. 39, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
```

```
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Pyridoxal phosphate biosynthetic protein pdxJ (PNP synthase).
GN   PDXJ OR BLL5064.
OS   Bradyrhizobium japonicum.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Bradyrhizobium.
OX   NCBI_TaxID=375;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=USDA 110spc4;
RA   Mueller P., Stingel D.;
RT   "Extended DNA sequencing in the upstream region of sipF in
RT   Bradyrhizobium japonicum.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=USDA 110;
RX   MEDLINE=22484998; PubMed=12597275;
RA   Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA   Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA   Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA   Tabata S.;
RT   "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT   Bradyrhizobium japonicum USDA110.";
RL   DNA Res. 9:189-197(2002).
RN   [3]
RP   SEQUENCE OF 67-250 FROM N.A.
RC   STRAIN=USDA 110spc4;
RX   MEDLINE=99086246; PubMed=9870699;
RA   Bairl A., Mueller P.;
RT   "A second gene for type I signal peptidase in Bradyrhizobium
RT   japonicum, sipF, is located near genes involved in RNA processing and
RT   cell division.";
RL   Mol. Gen. Genet. 260:346-356(1998).
CC   -!- FUNCTION: Catalyzes the condensation of 1-deoxy-D-xylulose-5-
CC       phosphate (DXP) and 1-amino-3-(phosphohydroxy)propan-2-one to form
CC       pyridoxine 5'-phosphate (PNP) (By similarity).
CC   -!- PATHWAY: De novo synthesis of pyridoxine (vitamin B6) and
CC       pyridoxal phosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the pdxJ family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF065159; AAD02936.2; -.
DR   EMBL; AP005953; BAC50329.1; -.
DR   HSSP; P24223; 1HO1.
DR   HAMAP; MF_00279; -; 1.
DR   InterPro; IPR004569; PdxJ.
DR   Pfam; PF03740; PdxJ; 1.
DR   ProDom; PD013657; PdxJ; 1.
DR   TIGRFAMs; TIGR00559; pdxJ; 1.
KW   Pyridoxine biosynthesis; Complete proteome.
FT   CONFLICT    185    186       KL -> NV (IN REF. 1 AND 3).
FT   CONFLICT    238    240       SMR -> NA (IN REF. 1 AND 3).
SQ   SEQUENCE   250 AA;  27033 MW;  47DBD839C0FA59CC CRC64;

  Query Match           89.7%;  Score 26;  DB 1;  Length 250;
  Best Local Similarity   83.3%;  Pred. No. 33;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              :|||||
Db        149 LRAPVI 154


RESULT 7
AGAY_ECOLI
ID   AGAY_ECOLI     STANDARD;      PRT;   286 AA.
AC   P42908;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Tagatose-1,6-bisphosphate aldolase agaY (EC 4.1.2.-) (TBPA).
GN   AGAY OR KBA OR B3137 OR C3894 OR Z4491 OR ECS4017.
OS   Escherichia coli,
OS   Escherichia coli O6, and
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562, 217992, 83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX   MEDLINE=22388234; PubMed=12471157;
RA   Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA   Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA   Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA   Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT   "Extensive mosaic structure revealed by the complete genome sequence
RT   of uropathogenic Escherichia coli.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
RN   [5]
RP   DISCUSSION OF SEQUENCE.
RX   MEDLINE=97086503; PubMed=8932697;
RA   Reizer J., Ramseier T.M., Reizer A., Charbit A., Saier M.H. Jr.;
RT   "Novel phosphotransferase genes revealed by bacterial genome
RT   sequencing: a gene cluster encoding a putative N-acetylgalactosamine
RT   metabolic pathway in Escherichia coli.";
RL   Microbiology 142:231-250(1996).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.45 ANGSTROMS).
RX   MEDLINE=22050650; PubMed=11940603;
RA   Hall D.R., Bond C.S., Leonard G.A., Watt C.I., Berry A., Hunter W.N.;
RT   "Structure of tagatose-1,6-bisphosphate aldolase. Insight into chiral
RT   discrimination, mechanism, and specificity of class II aldolases.";
RL   J. Biol. Chem. 277:22018-22024(2002).
```

```
CC   -!- FUNCTION: Catalyzes the reversible condensation of
CC       dihydroxyacetone phosphate with glyceraldehyde 3-phosphate to
CC       produce tagatose 1,6-bisphosphate.
CC   -!- CATALYTIC ACTIVITY: D-tagatose 1,6-bisphosphate 1,6-bisphosphate =
CC       glycerone-phosphate + D-glyceraldehyde 3-phosphate.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit.
CC   -!- PATHWAY: N-acetylgalactosamine utilization.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SIMILARITY: Belongs to class II fructose-bisphosphate aldolase
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U18997; AAA57940.1; -.
DR   EMBL; AE000395; AAC76171.1; -.
DR   EMBL; AE016767; AAN82335.1; -.
DR   EMBL; AE005542; AAG58269.1; -.
DR   EMBL; AP002564; BAB37440.1; -.
DR   PIR; A85976; A85976.
DR   PIR; A91131; A91131.
DR   PIR; E65103; E65103.
DR   PDB; 1GVF; 18-JUN-02.
DR   EcoGene; EG12768; agaY.
DR   InterPro; IPR000771; K_bP_aldolase.
DR   Pfam; PF01116; F_bP_aldolase; 1.
DR   TIGRFAMs; TIGR00167; cbbA; 1.
DR   PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR   PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
KW   Lyase; Zinc; Complete proteome; 3D-structure.
FT   METAL        83     83       ZINC.
FT   METAL       180    180       ZINC.
FT   METAL       208    208       ZINC.
SQ   SEQUENCE   286 AA;  31293 MW;  CAA42DF05C2B918B CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 286;
  Best Local Similarity   83.3%;  Pred. No. 38;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||:|||
Db         41 MRSPVI 46


RESULT 8
HEMZ_CHLPN
ID   HEMZ_CHLPN     STANDARD;      PRT;   327 AA.
AC   Q9Z7V1; Q9JSD1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ferrochelatase (EC 4.99.1.1) (Protoheme ferro-lyase) (Heme
DE   synthetase).
GN   HEMH OR CPN0603 OR CP0144 OR CPB0627.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TW-183;
RA   Geng M.M., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA   Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT   "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT   other Chlamydia strains based on whole genome sequence analysis.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Catalyzes the ferrous insertion into protoporphyrin IX.
CC   -!- CATALYTIC ACTIVITY: Protoporphyrin + Fe(2+) = protoheme + 2 H(+).
CC   -!- PATHWAY: Protoheme biosynthesis; last step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the ferrochelatase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001645; AAD18742.1; -.
DR   EMBL; AE002175; AAF38026.1; -.
DR   EMBL; AP002547; BAA98810.1; -.
DR   EMBL; AE017159; AAP98556.1; ALT_INIT.
DR   PIR; E72057; E72057.
DR   PIR; H86565; H86565.
DR   TIGR; CP0144; -.
DR   HAMAP; MF_00323; -; 1.
DR   InterPro; IPR001015; Ferrochelatase.
DR   Pfam; PF00762; Ferrochelatase; 1.
DR   ProDom; PD002792; Ferrochelatase; 1.
DR   TIGRFAMs; TIGR00109; hemH; 1.
DR   PROSITE; PS00534; FERROCHELATASE; FALSE_NEG.
KW   Porphyrin biosynthesis; Heme biosynthesis; Lyase; Iron;
KW   Complete proteome.
FT   METAL       187    187       IRON (BY SIMILARITY).
FT   METAL       265    265       IRON (BY SIMILARITY).
FT   CONFLICT    254    254       V -> A (IN REF. 3).
SQ   SEQUENCE   327 AA;  37502 MW;  608A8DC2B2139D44 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 327;
  Best Local Similarity   83.3%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              :|||||
Db         86 LRAPVI 91


RESULT 9
PMBA_ECOLI
ID   PMBA_ECOLI     STANDARD;      PRT;   450 AA.
AC   P24231;
```

```
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   PmbA protein (TldE protein).
GN   PMBA OR TLDE OR B4235 OR Z5845 OR ECS5212.
OS   Escherichia coli, and
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562, 83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=91186828; PubMed=2082149;
RA   Rodriguez-Sainz M.C., Hernandez-Chico C., Moreno F.;
RT   "Molecular characterization of pmbA, an Escherichia coli chromosomal
RT   gene required for the production of the antibiotic peptide MccB17.";
RL   Mol. Microbiol. 4:1921-1932(1990).
RN   [2]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   STRAIN=K12;
RX   MEDLINE=96177756; PubMed=8604133;
RA   Murayama N., Shimizu H., Takiguchi S., Baba Y., Amino H.,
RA   Horiuchi T., Sekimizu K., Miki T.;
RT   "Evidence for involvement of Escherichia coli genes pmbA, csrA and a
RT   previously unrecognized gene tldD, in the control of DNA gyrase by
RT   letD (ccdB) of sex factor F.";
RL   J. Mol. Biol. 256:483-502(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=95334362; PubMed=7610040;
RA   Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA   Blattner F.R.;
RT   "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT   region from 92.8 through 100 minutes.";
RL   Nucleic Acids Res. 23:2105-2119(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
CC   -!- FUNCTION: MAY FACILITATE THE SECRETION OF THE PEPTIDE ANTIBIOTIC
CC       MICROCIN B17 (MCCB17) BY COMPLETING ITS MATURATION. SUPPRESSES THE
CC       INHIBITORY ACTIVITY OF THE CARBON STORAGE REGULATOR (CSRA).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the tldD/pmbA family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X54152; CAA38091.1; -.
DR   EMBL; D44452; BAA07915.1; -.
DR   EMBL; U14003; AAA97132.1; -.
DR   EMBL; AE000494; AAC77192.1; -.
DR   EMBL; AE005655; AAG59432.1; -.
DR   EMBL; AP002568; BAB38635.1; -.
DR   PIR; D86121; D86121.
DR   PIR; D91280; D91280.
DR   PIR; S13730; S13730.
DR   EcoGene; EG10741; pmbA.
DR   InterPro; IPR002510; Peptidase_U62.
DR   Pfam; PF01523; PmbA_TldD; 1.
KW   Complete proteome.
SQ   SEQUENCE   450 AA;  48369 MW;  A062969197B52B5E CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 450;
  Best Local Similarity   83.3%;  Pred. No. 62;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |:||||
Db        242 MKAPVI 247


RESULT 10
CH10_XANAC
ID   CH10_XANAC     STANDARD;      PRT;    95 AA.
AC   Q8RIT8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   10 kDa chaperonin (Protein Cpn10) (groES protein).
GN   GROS OR GROES OR XAC0541 OR XCC0522.
OS   Xanthomonas axonopodis (pv. citri),
OS   Xanthomonas campestris (pv. campestris), and
OS   Xanthomonas campestris (pv. phaseoli).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=92829, 340, 29445;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=X.a.citri, and X.c.campestris;
RC   STRAIN=306 / ATCC 13902 / XV 101, and ATCC 33913 / NCPPB 528;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=X.c.phaseoli;
RA   Hsu C.-C., Yu Y.-J., Yang M.-T.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to Cpn60 in the presence of Mg-ATP and suppresses
CC       the ATPase activity of the latter.
CC   -!- SUBUNIT: Heptamer of 7 subunits arranged in a ring (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
```

```
CC   -!- SIMILARITY: Belongs to the groES chaperonin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE011680; AAM35430.1; -.
DR   EMBL; AE012149; AAM39838.1; -.
DR   EMBL; AF426387; AAL74149.1; -.
DR   HAMAP; MF_00580; -; 1.
DR   InterPro; IPR001476; Chaprnin_Cpn10.
DR   Pfam; PF00166; cpn10; 1.
DR   PRINTS; PR00297; CHAPERONIN10.
DR   ProDom; PD000566; Chaprnin_Cpn10; 1.
DR   PROSITE; PS00681; CHAPERONINS_CPN10; 1.
KW   Chaperone; Complete proteome.
SQ   SEQUENCE   95 AA;  9979 MW;  F0F0724BD153CBE3 CRC64;

  Query Match             86.2%;  Score 25;  DB 1;  Length 95;
  Best Local Similarity   66.7%;  Pred. No. 21;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              :||||:
Db         54 LRAPVV 59


RESULT 11
HIS5_STRAW
ID   HIS5_STRAW     STANDARD;      PRT;   215 AA.
AC   Q82AA2;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Imidazole glycerol phosphate synthase subunit hisH (EC 2.4.2.-) (IGP
DE   synthase glutamine amidotransferase subunit) (IGP synthase subunit
DE   hisH) (ImGP synthase subunit hisH) (IGPS subunit hisH).
GN   HISH OR SAV6157.
OS   Streptomyces avermitilis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=33903;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=21477403; PubMed=11572948;
RA   Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA   Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=22608306; PubMed=12692562;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industrial
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
CC   -!- FUNCTION: IGPS catalyzes the conversion of PRFAR and glutamine to
CC       IGP, AICAR and glutamate. The hisH subunit provides the glutamine
CC       amidotransferase activity that produces the ammonia necessary to
CC       hisF for the synthesis of IGP and AICAR (By similarity).
CC   -!- CATALYTIC ACTIVITY: 5-[(5-phospho-1-deoxyribulos-1-
CC       ylamino)methylideneamino]-1-(5-phosphoribosyl)imidazole-4-
CC       carboxamide + L-glutamine = imidazole-glycerol phosphate + 5-
CC       aminoimidazol-4-carboxamide ribonucleotide + L-glutamate + H(2)O.
CC   -!- PATHWAY: Histidine biosynthesis; fifth step.
CC   -!- SUBUNIT: Heterodimer of hisH and hisF (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP005045; BAC73868.1; -.
DR   HAMAP; MF_00278; -; 1.
DR   InterPro; IPR000991; GATase_1.
DR   Pfam; PF00117; GATase; 1.
DR   PROSITE; PS00442; GATASE_TYPE_I; 1.
KW   Histidine biosynthesis; Transferase; Glutamine amidotransferase;
KW   Complete proteome.
FT   ACT_SITE     86     86       BY SIMILARITY.
FT   ACT_SITE    196    196       BY SIMILARITY.
FT   ACT_SITE    198    198       BY SIMILARITY.
SQ   SEQUENCE   215 AA;  23052 MW;  DDAD5CC85943CBBE CRC64;

  Query Match             86.2%;  Score 25;  DB 1;  Length 215;
  Best Local Similarity   66.7%;  Pred. No. 51;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||::
Db        167 MRAPLV 172


RESULT 12
YFG0_YEAST
ID   YFG0_YEAST     STANDARD;      PRT;   222 AA.
AC   P43544;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Hypothetical UPF0030 protein YFL060C.
GN   YFL060C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=95400292; PubMed=7670463;
RA   Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA   Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA   Yamazaki M., Tashiro H., Eki T.;
RT   "Analysis of the nucleotide sequence of chromosome VI from
RT   Saccharomyces cerevisiae.";
RL   Nat. Genet. 10:261-268(1995).
CC   -!- SIMILARITY: Belongs to the UPF0030 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D50617; BAA09181.1; -.
DR   PIR; S56195; S56195.
DR   GermOnline; 140096; -.
DR   SGD; S0001834; YFL060C.
DR   GO; GO:0005515; F:protein binding; IPI.
```

```
DR   GO; GO:0008614; P:pyridoxine metabolism; IDA.
DR   GO; GO:0009228; P:thiamin biosynthesis; IMP.
DR   InterPro; IPR002161; SNO.
DR   Pfam; PF01174; SNO; 1.
DR   PROSITE; PS01236; UPF0030; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   222 AA;  25132 MW;  C7A0EFA9826FB77C CRC64;

  Query Match             86.2%;  Score 25;  DB 1;  Length 222;
  Best Local Similarity   83.3%;  Pred. No. 53;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              :|||||
Db        151 IRAPVI 156


RESULT 13
YN74_YEAST
ID   YN74_YEAST     STANDARD;      PRT;   222 AA.
AC   P53823;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Hypothetical UPF0030 protein YNL334C.
GN   YNL334C OR N0285.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97313269; PubMed=9169873;
RA   Philippsen P., Kleine K., Pohelmann R., Duesterhoeft A., Hamberg K.,
RA   Hegemann J.H., Obermaier B., Urrestarazu L.A., Aert R., Albermann K.,
RA   Altmann R., Andre B., Baladron V., Ballesta J.P.G., Becam A.-M.,
RA   Beinhauer J., Boskovic J., Buitrago M.J., Bussereau F., Coster F.,
RA   Crouzet M., D'Angelo M., Dal Pero F., De Antoni A., Del Rey F.,
RA   Doignon F., Domdey H., Dubois E., Fiedler T., Fleig U., Floeth M.,
RA   Fritz C., Gaillardin C., Garcia-Cantalejo J.M., Glansdorff N.,
RA   Goffeau A., Gueldener U., Herbert C.J., Heumann K., Heuss-Neitzel D.,
RA   Hilbert H., Hinni K., Iraqui Houssaini I., Jacquet M., Jimenez A.,
RA   Jonniaux J.-L., Karpfinger-Hartl L., Lanfranchi G., Lepingle A.,
RA   Levesque H., Lyck R., Maftahi M., Mallet L., Maurer K.C.T.,
RA   Messenguy F., Mewes H.-W., Moestl D., Nasr F., Nicaud J.-M.,
RA   Niedenthal R.K., Pandolfo D., Pierard A., Piravandi E., Planta R.J.,
RA   Pohl T.M., Purnelle B., Rebischung C., Remacha M., Revuelta J.L.,
RA   Rinke M., Saiz J.E., Sartorello F., Scherens B., Sen-Gupta M.,
RA   Soler-Mira A., Urbanus J.H M., Valle G., Van Dyck L., Verhasselt P.,
RA   Vierendeels F., Vissers S., Voet M., Volckaert G., Wach A.,
RA   Wambutt R., Wedler H., Zollner A., Hani J.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome XIV
RT   and its evolutionary implications.";
RL   Nature 387:93-98(1997).
CC   -!- SIMILARITY: Belongs to the UPF0030 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z71610; CAA96268.1; -.
DR   PIR; S63320; S63320.
DR   GermOnline; 143340; -.
DR   SGD; S0005278; YNL334C.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0008614; P:pyridoxine metabolism; IDA.
DR   GO; GO:0009228; P:thiamin biosynthesis; IMP.
DR   InterPro; IPR002161; SNO.
DR   Pfam; PF01174; SNO; 1.
DR   PROSITE; PS01236; UPF0030; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   222 AA;  25208 MW;  6755EE1D826CAABA CRC64;

  Query Match             86.2%;  Score 25;  DB 1;  Length 222;
  Best Local Similarity   83.3%;  Pred. No. 53;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              :|||||
Db        151 IRAPVI 156


RESULT 14
YMY5_YEAST
ID   YMY5_YEAST     STANDARD;      PRT;   224 AA.
AC   Q03144;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Hypothetical UPF0030 protein YMR095C.
GN   YMR095C OR YM6543.02C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=97313268; PubMed=9169872;
RA   Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA   Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA   Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA   Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT   XIII.";
RL   Nature 387:90-93(1997).
CC   -!- SIMILARITY: Belongs to the UPF0030 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z49807; CAA89896.1; -.
DR   PIR; S55081; S55081.
DR   GermOnline; 142763; -.
DR   SGD; S0004701; YMR095C.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0008614; P:pyridoxine metabolism; IDA.
DR   InterPro; IPR002161; SNO.
DR   Pfam; PF01174; SNO; 1.
DR   PROSITE; PS01236; UPF0030; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   224 AA;  24906 MW;  C0EC1B2382581D40 CRC64;

  Query Match             86.2%;  Score 25;  DB 1;  Length 224;
  Best Local Similarity   83.3%;  Pred. No. 53;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              :|||||
Db        160 IRAPVI 165


RESULT 15
HIS6_RHIME
ID   HIS6_RHIME     STANDARD;      PRT;   258 AA.
AC   Q92TB3;
DT   28-FEB-2003 (Rel. 41, Created)
```

```
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Imidazole glycerol phosphate synthase subunit hisF (EC 4.1.3.-) (IGP
DE   synthase cyclase subunit) (IGP synthase subunit hisF) (ImGP synthase
DE   subunit hisF) (IGPS subunit hisF).
GN   HISF OR R00051 OR SMC02569.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396507; PubMed=11481430;
RA   Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA   Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA   Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA   Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA   Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT   "Analysis of the chromosome sequence of the legume symbiont
RT   Sinorhizobium meliloti strain 1021.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC   -!- FUNCTION: IGPS catalyzes the conversion of PRFAR and glutamine to
CC       IGP, AICAR and glutamate. The hisF subunit catalyzes the
CC       cyclization activity that produces IGP and AICAR from PRFAR using
CC       the ammonia provided by the hisH subunit (By similarity).
CC   -!- CATALYTIC ACTIVITY: 5-[(5-phospho-1-deoxyribulos-1-
CC       ylamino)methylideneamino]-1-(5-phosphoribosyl)imidazole-4-
CC       carboxamide + L-glutamine = imidazole-glycerol phosphate + 5-
CC       aminoimidazol-4-carboxamide ribonucleotide + L-glutamate + H(2)O.
CC   -!- PATHWAY: Histidine biosynthesis; fifth step.
CC   -!- SUBUNIT: Heterodimer of hisH and hisF (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the hisA / hisF family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL591782; CAC41438.1; -.
DR   HAMAP; MF_01013; -; 1.
DR   InterPro; IPR003009; FMN_enzyme.
DR   InterPro; IPR006062; His_biosynth.
DR   InterPro; IPR004651; HisF.
DR   Pfam; PF00977; His_biosynth; 1.
DR   TIGRFAMs; TIGR00735; hisF; 1.
KW   Histidine biosynthesis; Lyase; Complete proteome.
FT   ACT_SITE     12     12       POTENTIAL.
FT   ACT_SITE    131    131       POTENTIAL.
SQ   SEQUENCE   258 AA;  27312 MW;  56125BFBF22F39D3 CRC64;

  Query Match             86.2%;  Score 25;  DB 1;  Length 258;
  Best Local Similarity   83.3%;  Pred. No. 62;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              :|||||
Db        200 VRAPVI 205


Search completed: September 18, 2004, 03:43:44
Job time : 11.7857 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 03:24:30 ; Search time 31.2857 Seconds
(without alignments)
60.510 Million cell updates/sec

Title: US-10-029-926B-8
Perfect score: 29
Sequence: 1 MRAPVI 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	249	16	Q9A2W4	Q9a2w4 caulobacter
2	29	100.0	442	2	Q53093	Q53093 rhodobacter
3	29	100.0	442	2	Q53153	Q53153 rhodobacter
4	29	100.0	629	5	O61371	O61371 caenorhabdi
5	29	100.0	947	11	Q80TL6	Q80tl6 mus musculu
6	28	96.6	376	16	Q8P9F5	Q8p9f5 xanthomonas
7	28	96.6	541	16	Q8NT69	Q8nt69 corynebacte
8	28	96.6	713	5	Q7YYY3	Q7yyy3 cryptospori
9	27	93.1	247	17	O27295	O27295 methanobact
10	27	93.1	376	2	Q9L778	Q9l778 xanthomonas
11	27	93.1	376	16	Q8PL69	Q8pl69 xanthomonas
12	27	93.1	416	16	Q9ACQ9	Q9acq9 streptomyce
13	27	93.1	449	16	Q89X81	Q89x81 bradyrhizob
14	27	93.1	771	13	Q98TH2	Q98th2 brachydanio
15	26	89.7	162	5	Q8ILB7	Q8ilb7 plasmodium
16	26	89.7	286	2	Q9KIP8	Q9kip8 escherichia
17	26	89.7	353	8	Q9TE71	Q9te71 cylindrothe
18	26	89.7	353	8	Q9TE68	Q9te68 nitzschia f
19	26	89.7	450	16	Q8FAF2	Q8faf2 escherichia
20	26	89.7	450	16	Q8XGJ7	Q8xgj7 salmonella
21	26	89.7	450	16	Q83P54	Q83p54 shigella fl
22	26	89.7	450	16	Q7UAN7	Q7uan7 shigella fl
23	26	89.7	463	16	Q7UKY8	Q7uky8 rhodopirell
24	26	89.7	535	16	Q8ERE8	Q8ere8 oceanobacil
25	26	89.7	583	5	Q9VKE5	Q9vke5 drosophila
26	26	89.7	1127	13	Q7T1R3	Q7t1r3 brachydanio
27	26	89.7	1449	5	Q9V917	Q9v917 drosophila
28	25	86.2	93	12	Q8QNH1	Q8qnh1 ectocarpus
29	25	86.2	94	2	P97101	P97101 alcaligenes
30	25	86.2	103	16	Q839I3	Q839i3 enterococcu
31	25	86.2	110	10	Q9LJ06	Q9lj06 oryza sativ
32	25	86.2	119	10	Q7XMD2	Q7xmd2 oryza sativ
33	25	86.2	138	5	Q22499	Q22499 caenorhabdi
34	25	86.2	147	2	Q9X699	Q9x699 rhizobium m
35	25	86.2	147	16	Q92W24	Q92w24 rhizobium m
36	25	86.2	147	16	Q92VN7	Q92vn7 rhizobium m
37	25	86.2	149	16	Q9WXT6	Q9wxt6 thermotoga
38	25	86.2	154	16	Q9PG98	Q9pg98 xylella fas
39	25	86.2	154	16	Q87AZ3	Q87az3 xylella fas
40	25	86.2	159	16	Q893T7	Q893t7 clostridium
41	25	86.2	163	16	Q9A8P1	Q9a8p1 caulobacter
42	25	86.2	182	16	Q8EDN5	Q8edn5 shewanella
43	25	86.2	188	1	Q9UWY2	Q9uwy2 sulfolobus
44	25	86.2	188	16	Q7UWU0	Q7uwu0 rhodopirell
45	25	86.2	195	16	Q9RUG7	Q9rug7 deinococcus

ALIGNMENTS

RESULT 1
Q9A2W4
ID Q9A2W4 PRELIMINARY; PRT; 249 AA.
AC Q9A2W4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hydrolase, alpha/beta hydrolase fold family.
GN CC3442.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -!- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.
DR EMBL; AE006003; AAK25404.1; -.
DR PIR; H87675; H87675.
DR TIGR; CC3442; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR InterPro; IPR000073; A/b_hydrolase.
DR InterPro; IPR003089; AB_hydrolase.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00561; abhydrolase; 1.
DR PRINTS; PR00111; ABHYDROLASE.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 249 AA; 26772 MW; 491350127DD300AE CRC64;

```
  Query Match             100.0%;  Score 29;  DB 16;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 84;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db          1 MRAPVI 6


RESULT 2
Q53093
ID   Q53093      PRELIMINARY;      PRT;   442 AA.
AC   Q53093;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   FliI.
GN   FLII.
OS   Rhodobacter sphaeroides (Rhodopseudomonas sphaeroides).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC   Rhodobacteraceae; Rhodobacter.
OX   NCBI_TaxID=1063;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2.4.1;
RX   MEDLINE=96200857; PubMed=8621091;
RA   Ballado T., Campos A., Camarena L., Dreyfus G.;
RT   "Flagellar genes from Rhodobacter sphaeroides are homologous to genes
RT   of the fliF operon of Salmonella typhimurium and to the type-III
RT   secretion system.";
RL   Gene 170:69-72(1996).
CC   -!- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
DR   EMBL; U31090; AAB07344.1; -.
DR   PIR; JC4733; JC4733.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0016887; F:ATPase activity; IEA.
DR   GO; GO:0015078; F:hydrogen ion transporter activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0000166; F:nucleotide binding; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   GO; GO:0006811; P:ion transport; IEA.
DR   GO; GO:0015031; P:protein transport; IEA.
DR   GO; GO:0015992; P:proton transport; IEA.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR000194; ATPase_a/bcentre.
DR   InterPro; IPR005714; FliI_YscN.
DR   Pfam; PF00006; ATP-synt_ab; 1.
DR   SMART; SM00382; AAA; 1.
DR   TIGRFAMs; TIGR01026; fliI_yscN; 1.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
KW   ATP-binding; Hydrogen ion transport; Hydrolase; Ion transport;
KW   Transport.
SQ   SEQUENCE   442 AA;  46852 MW;  E6D35531F5A59BAE CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 442;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db         77 MRAPVI 82


RESULT 3
Q53153
ID   Q53153      PRELIMINARY;      PRT;   442 AA.
AC   Q53153;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   FliI protein.
GN   FLII.
OS   Rhodobacter sphaeroides (Rhodopseudomonas sphaeroides).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC   Rhodobacteraceae; Rhodobacter.
OX   NCBI_TaxID=1063;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WS8;
RX   MEDLINE=96327148; PubMed=8759796;
RA   Goodfellow I.G., Pollitt C.E., Sockett R.E.;
RT   "Cloning of the fliI gene from Rhodobacter sphaeroides WS8 by analysis
RT   of a transposon mutant with impaired motility.";
RL   FEMS Microbiol. Lett. 142:111-116(1996).
CC   -!- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
DR   EMBL; X97201; CAA65834.1; -.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0016887; F:ATPase activity; IEA.
DR   GO; GO:0015078; F:hydrogen ion transporter activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0000166; F:nucleotide binding; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   GO; GO:0006811; P:ion transport; IEA.
DR   GO; GO:0015031; P:protein transport; IEA.
DR   GO; GO:0015992; P:proton transport; IEA.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR000194; ATPase_a/bcentre.
DR   InterPro; IPR005714; FliI_YscN.
DR   Pfam; PF00006; ATP-synt_ab; 1.
DR   SMART; SM00382; AAA; 1.
DR   TIGRFAMs; TIGR01026; fliI_yscN; 1.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
KW   ATP-binding; Hydrogen ion transport; Hydrolase; Ion transport;
KW   Transport.
SQ   SEQUENCE   442 AA;  46813 MW;  AD070D4E17FD3CC3 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 442;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db         77 MRAPVI 82


RESULT 4
O61371
ID   O61371      PRELIMINARY;      PRT;   629 AA.
AC   O61371; Q9TXY7;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Acetylcholinesterase (EC 3.1.1.7) (Hypothetical protein).
GN   ACE-2 OR Y44E3A.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RX   MEDLINE=98198570; PubMed=9539167;
RA   Grauso M., Culetto E., Combes D., Fedon Y., Toutant J.P., Arpagaus M.;
RT   "Existence of four acetylcholinesterase genes in the nematodes
RT   Caenorhabditis elegans and Caenorhabditis briggsae.";
RL   FEBS Lett. 424:279-284(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RA   Grauso M., Culetto E., Fedon Y., Combes D., Toutant J.P., Arpagaus M.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
```

```
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Woessner J., Graves T., Keppler D.;
RT   "The sequence of C. elegans cosmid Y44E3A.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
DR   EMBL; AF025378; AAC14016.2; -.
DR   EMBL; AF106589; AAC78228.2; -.
DR   PIR; T33842; T33842.
DR   PIR; T37255; T37255.
DR   HSSP; P21836; 1MAA.
DR   WormPep; Y44E3A.2; CE28363.
DR   GO; GO:0003990; F:acetylcholinesterase activity; IEA.
DR   GO; GO:0004104; F:cholinesterase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   InterPro; IPR002018; CarbesteraseB.
DR   InterPro; IPR000997; Cholinesterase.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF00135; COesterase; 1.
DR   PRINTS; PR00878; CHOLNESTRASE.
DR   PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
KW   Hydrolase; Hypothetical protein.
SQ   SEQUENCE   629 AA;  70863 MW;  74940F512FEDF869 CRC64;

  Query Match             100.0%;  Score 29;  DB 5;  Length 629;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db          1 MRAPVI 6


RESULT 5
Q80TL6
ID   Q80TL6      PRELIMINARY;      PRT;   947 AA.
AC   Q80TL6;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   MKIAA1042 protein (Fragment).
GN   MKIAA1042.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22579291; PubMed=12693553;
RA   Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA   Nakajima D., Nagase T., Ohara O., Koga H.;
RT   "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT   II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT   cDNAs identified by screening of terminal sequences of cDNA clones
RT   randomly sampled from size-fractionated libraries.";
RL   DNA Res. 10:35-48(2003).
DR   EMBL; AK122426; BAC65708.1; -.
DR   GO; GO:0005634; C:nucleus; ISS.
DR   GO; GO:0005886; C:plasma membrane; ISS.
DR   GO; GO:0005515; F:protein binding; ISS.
DR   GO; GO:0005102; F:receptor binding; ISS.
DR   GO; GO:0006493; P:O-linked glycosylation; ISS.
DR   GO; GO:0006605; P:protein targeting; ISS.
DR   GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; ISS.
DR   InterPro; IPR006933; HAP1_N.
DR   InterPro; IPR000585; Hemopexin.
DR   Pfam; PF04849; HAP1_N; 1.
DR   PROSITE; PS00024; HEMOPEXIN; 1.
FT   NON_TER      1      1
SQ   SEQUENCE   947 AA;  105486 MW;  0987284C6ACF23A5 CRC64;

  Query Match             100.0%;  Score 29;  DB 11;  Length 947;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||||
Db        924 MRAPVI 929


RESULT 6
Q8P9F5
ID   Q8P9F5      PRELIMINARY;      PRT;   376 AA.
AC   Q8P9F5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Flagellar protein.
GN   FLHB OR XCC1910.
OS   Xanthomonas campestris (pv. campestris).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=340;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33913 / NCPPB 528;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
DR   EMBL; AE012294; AAM41199.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0009306; P:protein secretion; IEA.
DR   InterPro; IPR006135; Bac_Export_2.
DR   InterPro; IPR006136; FlhB.
DR   Pfam; PF01312; Bac_export_2; 1.
DR   PRINTS; PR00950; TYPE3IMSPROT.
DR   TIGRFAMs; TIGR00328; flhB; 1.
KW   Complete proteome.
SQ   SEQUENCE   376 AA;  41490 MW;  406778385A158910 CRC64;

  Query Match             96.6%;  Score 28;  DB 16;  Length 376;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |||||:
Db        285 MRAPVV 290
```

```
RESULT 7
Q8NT69
ID   Q8NT69      PRELIMINARY;      PRT;   541 AA.
AC   Q8NT69;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   ResB protein required for cytochrome c biosynthesis.
GN   CGL0441.
OS   Corynebacterium glutamicum (Brevibacterium flavum).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=1718;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA   Nakagawa S.;
RT   "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005275; BAB97834.1; -.
DR   InterPro; IPR007816; ResB.
DR   Pfam; PF05140; ResB; 1.
KW   Complete proteome.
SQ   SEQUENCE   541 AA;  61244 MW;  37F42BA0A74F78BB CRC64;

  Query Match             96.6%;  Score 28;  DB 16;  Length 541;
  Best Local Similarity   83.3%;  Pred. No. 3.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |||||:
Db        103 MRAPVV 108


RESULT 8
Q7YYY3
ID   Q7YYY3      PRELIMINARY;      PRT;   713 AA.
AC   Q7YYY3;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   1MB.818.
OS   Cryptosporidium parvum.
OC   Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC   Cryptosporidiidae; Cryptosporidium.
OX   NCBI_TaxID=5807;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Iowa;
RA   Bankier A.T., Spriggs H.F., Fartmann B., Konfortov B.A., Madera M.,
RA   Vogel C., Teichmann S.A., Ivens A., Dear P.H.;
RT   "Integrated mapping, chromosomal sequencing and sequence analysis of
RT   Cryptosporidium parvum.";
RL   Genome Res. 0:0-0(2003).
DR   EMBL; BX538353; CAD98347.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   713 AA;  82980 MW;  101EE54E3C2163CA CRC64;

  Query Match             96.6%;  Score 28;  DB 5;  Length 713;
  Best Local Similarity   83.3%;  Pred. No. 4.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||:|
Db        106 MRAPII 111


RESULT 9
O27295
ID   O27295      PRELIMINARY;      PRT;   247 AA.
AC   O27295;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Coenzyme PQQ synthesis protein III.
GN   MTH1227.
OS   Methanobacterium thermoautotrophicum.
OC   Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC   Methanobacteriaceae; Methanothermobacter.
OX   NCBI_TaxID=187420;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Delta H;
RX   MEDLINE=98037514; PubMed=9371463;
RA   Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
RA   Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA   Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA   Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA   Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA   McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA   Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   deltaH: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
DR   EMBL; AE000890; AAB85716.1; -.
DR   PIR; H69030; H69030.
DR   InterPro; IPR007197; Radical_SAM.
DR   Pfam; PF04055; Radical_SAM; 1.
KW   Complete proteome.
SQ   SEQUENCE   247 AA;  27374 MW;  DD1F0C0251DC5217 CRC64;

  Query Match             93.1%;  Score 27;  DB 17;  Length 247;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||::
Db          4 MRAPIV 9


RESULT 10
Q9L778
ID   Q9L778      PRELIMINARY;      PRT;   376 AA.
AC   Q9L778;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Flagellar biosynthetic protein FlhB.
GN   FLHB.
OS   Xanthomonas oryzae (pv. oryzae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=64187;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21071238; PubMed=11204784;
RA   Shen Y., Chern M.-S., Silva F.G., Ronald P.;
RT   "Isolation of a Xanthomonas oryzae pv. oryzae flagellar operon region
RT   and molecular characterization of flhF.";
RL   Mol. Plant Microbe Interact. 14:204-213(2001).
DR   EMBL; AF226282; AAF33828.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0009306; P:protein secretion; IEA.
DR   InterPro; IPR006135; Bac_Export_2.
DR   InterPro; IPR006136; FlhB.
DR   Pfam; PF01312; Bac_export_2; 1.
DR   PRINTS; PR00950; TYPE3IMSPROT.
DR   TIGRFAMs; TIGR00328; flhB; 1.
SQ   SEQUENCE   376 AA;  41154 MW;  68514424C3D1E452 CRC64;

  Query Match             93.1%;  Score 27;  DB 2;  Length 376;
  Best Local Similarity   66.7%;  Pred. No. 3.8e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 MRAPVI 6
              ||||::
Db        285 MRAPIV 290


RESULT 11
Q8PL69
ID   Q8PL69      PRELIMINARY;      PRT;   376 AA.
AC   Q8PL69;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Flagellar protein.
GN   FLHB OR XAC1937.
OS   Xanthomonas axonopodis (pv. citri).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=92829;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=306 / ATCC 13902 / XV 101;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
DR   EMBL; AE011829; AAM36799.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0009306; P:protein secretion; IEA.
DR   InterPro; IPR006135; Bac_Export_2.
DR   InterPro; IPR006136; FlhB.
DR   Pfam; PF01312; Bac_export_2; 1.
DR   PRINTS; PR00950; TYPE3IMSPROT.
DR   TIGRFAMs; TIGR00328; flhB; 1.
KW   Complete proteome.
SQ   SEQUENCE   376 AA;  41333 MW;  615AE986871936AC CRC64;

  Query Match             93.1%;  Score 27;  DB 16;  Length 376;
  Best Local Similarity   66.7%;  Pred. No. 3.8e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||::
Db        285 MRAPIV 290


RESULT 12
Q9ACQ9
ID   Q9ACQ9      PRELIMINARY;      PRT;   416 AA.
AC   Q9ACQ9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative transposon transposase.
GN   SCP1.259.
OS   Streptomyces coelicolor.
OG   Plasmid SCP1.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Brown S.P., Murphy L.D., Harris D.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=98241550; PubMed=9573173;
RA   Redenbach M., Ikeda K., Yamasaki M., Kinashi H.;
RT   "Cloning and physical mapping of the EcoRI fragments of the giant
RT   linear plasmid SCP1.";
RL   J. Bacteriol. 180:2796-2799(1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL590464; CAC36785.1; -.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR   InterPro; IPR001207; Trnspsase_mutatr.
DR   Pfam; PF00872; Transpo_mutator; 1.
DR   PROSITE; PS01007; TRANSPOSASE_MUTATOR; 1.
KW   Plasmid; Complete proteome.
SQ   SEQUENCE   416 AA;  45815 MW;  9CBFDD2D33F4B1F2 CRC64;

  Query Match             93.1%;  Score 27;  DB 16;  Length 416;
  Best Local Similarity   83.3%;  Pred. No. 4.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |||||:
Db        232 MRAPVL 237


RESULT 13
Q89X81
ID   Q89X81      PRELIMINARY;      PRT;   449 AA.
AC   Q89X81;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Blr0433 protein.
GN   BLR0433.
OS   Bradyrhizobium japonicum.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Bradyrhizobium.
OX   NCBI_TaxID=375;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=USDA 110;
RX   MEDLINE=22484998; PubMed=12597275;
RA   Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA   Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA   Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA   Tabata S.;
```

```
RT   "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT   Bradyrhizobium japonicum USDA110.";
RL   DNA Res. 9:189-197(2002).
DR   EMBL; AP005936; BAC45698.1; -.
DR   GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002886; Peptidase_M37.
DR   Pfam; PF01551; Peptidase_M37; 1.
KW   Complete proteome.
SQ   SEQUENCE   449 AA;  47625 MW;  0374885B47AB3511 CRC64;

  Query Match           93.1%;  Score 27;  DB 16;  Length 449;
  Best Local Similarity  83.3%;  Pred. No. 4.6e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |||||:
Db          1 MRAPVL 6


RESULT 14
Q98TH2
ID   Q98TH2      PRELIMINARY;      PRT;   771 AA.
AC   Q98TH2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Groucho-related gene 3 protein.
GN   GRO3 OR GRG3.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21294966; PubMed=11401394;
RA   Kobayashi M., Nishikawa K., Suzuki T., Yamamoto M.;
RT   "The homeobox protein Six3 interacts with the Groucho corepressor and
RT   acts as a transcriptional repressor in eye and forebrain formation.";
RL   Dev. Biol. 232:315-326(2001).
DR   EMBL; AB040707; BAB40698.1; -.
DR   ZFIN; ZDB-GENE-010410-1; gro3.
DR   InterPro; IPR005617; TLE_N.
DR   InterPro; IPR001680; WD40.
DR   Pfam; PF03920; TLE_N; 1.
DR   Pfam; PF00400; WD40; 6.
DR   ProDom; PD000018; WD40; 1.
DR   SMART; SM00320; WD40; 7.
DR   PROSITE; PS00678; WD_REPEATS_1; 2.
DR   PROSITE; PS50082; WD_REPEATS_2; 2.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 2.
KW   Repeat; WD repeat.
SQ   SEQUENCE   771 AA;  83789 MW;  3B9CEB203826E458 CRC64;

  Query Match           93.1%;  Score 27;  DB 13;  Length 771;
  Best Local Similarity  83.3%;  Pred. No. 7.8e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              |||||:
Db        428 MRAPVL 433


RESULT 15
Q8ILB7
ID   Q8ILB7      PRELIMINARY;      PRT;   162 AA.
AC   Q8ILB7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Mitochondrial import inner membrane translocase subunit tim17,
DE   putative.
GN   PF14_0328.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3D7;
RX   MEDLINE=22255705; PubMed=12368864;
RA   Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA   Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA   Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA   Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA   Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA   Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA   McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA   Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA   Fraser C.M., Barrell B.;
RT   "Genome sequence of the human malaria parasite Plasmodium
RT   falciparum.";
RL   Nature 419:498-511(2002).
DR   EMBL; AE014821; AAN36941.1; -.
DR   GO; GO:0005744; C:mitochondrial inner membrane pre-sequence t. . .; IEA.
DR   GO; GO:0015450; F:protein translocase activity; IEA.
DR   GO; GO:0015031; P:protein transport; IEA.
DR   InterPro; IPR003397; Tim17_Tim22.
DR   Pfam; PF02466; Tim17; 1.
SQ   SEQUENCE   162 AA;  17909 MW;  F536704A462221FF CRC64;

  Query Match           89.7%;  Score 26;  DB 5;  Length 162;
  Best Local Similarity  66.7%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRAPVI 6
              ||||::
Db         58 MRAPIL 63


Search completed: September 18, 2004, 03:46:23
Job time : 35.2857 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 02:28:24 ; Search time 35.3571 Seconds
(without alignments)
39.956 Million cell updates/sec

Title: US-10-029-926B-114
Perfect score: 28
Sequence: 1 DYGMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	100.0	5	3	AAY95191	Aay95191 Anti-plat
2	28	100.0	5	5	ABG78239	Abg78239 Human Fv
3	28	100.0	5	5	ABG91930	Abg91930 Human ant
4	28	100.0	5	6	ABU11271	Abu11271 Human TAN
5	28	100.0	16	2	AAR49460	Aar49460 Factor VI
6	28	100.0	20	5	ABJ05002	Abj05002 A3 peptid
7	28	100.0	26	1	AAP61435	Aap61435 Factor VI
8	28	100.0	29	1	AAP50314	Aap50314 Peptide e
9	28	100.0	33	1	AAP50318	Aap50318 Peptide e
10	28	100.0	41	1	AAP50313	Aap50313 Thrombin
11	28	100.0	60	5	ABJ04948	Abj04948 A3 domain
12	28	100.0	73	3	AAB07206	Aab07206 Human Fac
13	28	100.0	98	3	AAB40073	Aab40073 Anti-hIL1
14	28	100.0	98	5	ABG78186	Abg78186 Human Fv
15	28	100.0	98	5	ABG91877	Abg91877 Human ant
16	28	100.0	98	6	ABO27087	Abo27087 Human ger
17	28	100.0	113	3	AAY95177	Aay95177 Anti-plat
18	28	100.0	113	3	AAY95178	Aay95178 Anti-plat
19	28	100.0	115	3	AAY95189	Aay95189 Anti-plat
20	28	100.0	115	3	AAY95190	Aay95190 Anti-plat
21	28	100.0	116	2	AAW19880	CEA-speci
22	28	100.0	117	2	AAR66312	Aar66312 Human imm
23	28	100.0	118	4	AAU02560	Aau02560 Anti-adip
24	28	100.0	120	2	AAR25204	Aar25204 OP-G2 mon
25	28	100.0	207	5	AAU98019	Aau98019 Human ace
26	28	100.0	211	1	AAP50091	Aap50091 Truncated
27	28	100.0	211	3	AAB07203	Aab07203 Human Fac
28	28	100.0	214	3	AAG38412	Aag38412 Arabidops
29	28	100.0	236	3	AAG29486	Aag29486 Arabidops
30	28	100.0	238	3	AAY95198	Aay95198 Anti-plat
31	28	100.0	239	5	ABP46004	Abp46004 Human BLy
32	28	100.0	239	5	ABP46027	Abp46027 Human BLy
33	28	100.0	239	5	ABP44926	Abp44926 Human BLy
34	28	100.0	244	6	AAO31136	Aao31136 Human CM0
35	28	100.0	246	5	ABG78329	Abg78329 Human Fv
36	28	100.0	246	5	ABG92026	Abg92026 Antibody
37	28	100.0	249	5	ABP30235	Abp30235 Streptoco
38	28	100.0	251	5	ABP45531	Abp45531 Human BLy
39	28	100.0	256	5	ABG78334	Abg78334 Human Fv
40	28	100.0	256	5	ABG92025	Abg92025 Antibody
41	28	100.0	258	7	ADC94375	Adc94375 E. faeciu
42	28	100.0	259	3	AAG32033	Aag32033 Arabidops
43	28	100.0	264	7	ADC97186	Adc97186 E. faeciu
44	28	100.0	266	5	ABG92020	Abg92020 Human ant
45	28	100.0	267	3	AAG29485	Aag29485 Arabidops

ALIGNMENTS

RESULT 1
AAY95191
ID AAY95191 standard; peptide; 5 AA.
XX
AC AAY95191;
XX
DT 29-AUG-2000 (first entry)
XX
DE Anti-platelet glycoprotein Ib human HIb-1 VH CDR1.
XX
KW Variable heavy chain; single chain antibody; scFv; human; HIb-1;
KW glycoprotein Ib alpha; platelet; aggregation; antiaggregant;
KW antithrombotic; thrombus; therapy; diagnostic; CDR1;
KW complementarity determining region.
XX
OS Homo sapiens.
XX
PN WO200026667-A1.
XX
PD 11-MAY-2000.
XX
PF 29-OCT-1999; 99WO-US025495.
XX
PR 30-OCT-1998; 98US-0106275P.
XX
PA (MILL/) MILLER J L.
XX
PI Miller JL;
XX
DR WPI; 2000-365744/31.
XX
PT Isolated nucleic acid molecule encoding anti-human platelet glycoprotein
PT Ib alpha molecule useful for producing antibodies which inhibit platelet
PT aggregation.
XX
PS Claim 14; Fig 5; 89pp; English.
XX
CC The present sequence is that of complementarity determining region 1
CC (CDR1) of the heavy chain variable region (VH) of human single chain
CC antibody (scFv) HIb-1 (see AAY95198), which is directed against platelet
CC glycoprotein Ib (GPIb). The HIb series of scFv was isolated from a human
CC synthetic VH and VL scFv library on the basis of their binding to
CC platelet GPIb. Whether displayed as surface proteins on a phagemid or
CC secreted as free scFv by Escherichia coli, the HIb scFv clones are
CC capable of inhibiting von Willebrand factor-dependent aggregation of
CC platelets. The scFv are composed of native human protein sequences and
CC are therefore attractive potential reagents for therapeutic purposes.
CC They provide a new class of antithrombotic agents, useful for the

```
CC   prevention of platelet-dependent thrombi in diseased arteries, bypass
CC   grafts, dialysis etc., and can also be used as diagnostic reagents.
CC   Methods of inhibiting aggregation of platelets, of binding human platelet
CC   GPIb alpha and of selecting a VH or VL region of an antibody that
CC   inhibits platelet aggregation are claimed. Fragments of the scFv VH or VL
CC   chain, including CDR fragments, are also claimed
XX
SQ   Sequence 5 AA;

  Query Match             100.0%;  Score 28;  DB 3;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          1 DYGMS 5


RESULT 2
ABG78239
ID   ABG78239 standard; peptide; 5 AA.
XX
AC   ABG78239;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Human Fv molecule hypervariable region related peptide #114.
XX
KW   Human; Fv molecule; hypervariable region; single chain Fv; cytostatic;
KW   disulfide Fv; dsFv; scFv; cancer; carcinoma; sarcoma; leukaemia; adenoma;
KW   lymphoma; myeloma; blastoma; seminoma; melanoma; acute myeloid leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200259264-A2.
XX
PD   01-AUG-2002.
XX
PF   31-DEC-2001; 2001WO-US049440.
XX
PR   29-DEC-2000; 2000US-00751181.
XX
PA   (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI   Hagai Y,  Lazarovits J,  Guy R,  Lipschitz O,  Szanton E,  Levanon A;
PI   Plaksin D,  Peretz T;
XX
DR   WPI; 2002-619166/66.
XX
PT   Novel peptide/polypeptide for cancer therapy has Fv molecule, construct
PT   or fragment, or construct of fragment with enhanced binding
PT   characteristics so as to selectively bind target cell in favor of other
PT   cells.
XX
PS   Claim 20; Page 208; 232pp; English.
XX
CC   The invention relates to a peptide or polypeptide comprising an Fv
CC   molecule, a construct or fragments or a construct of a fragment with
CC   enhanced binding characteristics which selectively and/or specifically
CC   binds to a target cell in favour of other cells, where binding is
CC   primarily determined by a first hypervariable region and Fv is a single
CC   chain Fv (scFv) or a disulfide Fv (dsFv). The peptide, optionally in
CC   association with or attached, coupled, combined, linked or fused to a
CC   pharmaceutical agent, is useful in the manufacture of a medicament, where
CC   the medicament has activity against a diseased cell, preferably a cancer
CC   cell (selected from carcinoma, sarcoma, leukaemia, adenoma, lymphoma,
CC   myeloma, blastoma, seminoma, and melanoma, where the leukaemia cell is an
CC   acute myeloid leukaemia cell). The peptide is also useful for preparing a
CC   composition for use in inhibiting the growth of a diseased or cancer
CC   cell. This sequence represents a human Fv molecule hypervariable region
CC   related peptide of the invention
XX
SQ   Sequence 5 AA;

  Query Match             100.0%;  Score 28;  DB 5;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          1 DYGMS 5


RESULT 3
ABG91930
ID   ABG91930 standard; peptide; 5 AA.
XX
AC   ABG91930;
XX
DT   04-DEC-2002  (first entry)
XX
DE   Human antibody fragment #114.
XX
KW   Human; antibody; epitope; cancer; tumour; cell rolling; inflammation;
KW   metastasis; hypervariable region; autoimmune disease; thrombosis;
KW   restenosis; leukaemia; inflammatory disease; cardiovascular disease;
KW   myocardial infarction; retinopathic disease; abnormal platelet function;
KW   sulphated tyrosine-dependent protein-protein interaction.
XX
OS   Homo sapiens.
XX
PN   WO200253700-A2.
XX
PD   11-JUL-2002.
XX
PF   31-DEC-2001; 2001WO-US049442.
XX
PR   29-DEC-2000; 2000US-00751181.
PR   29-DEC-2000; 2000US-0258948P.
XX
PA   (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI   Lazarovits J,  Hagai Y,  Plaksin D,  Vogel T,  Nimrod A,  Mar-Haim H;
PI   Szanthon E,  Richter T,  Amit B,  Kooperman L,  Peretz T,  Levanon A;
XX
DR   WPI; 2002-674776/72.
XX
PT   Novel isolated epitope present on cancer cells and important in
PT   physiological phenomena such as cell rolling, metastasis and
PT   inflammation, for treating autoimmune, inflammatory or cardiovascular
PT   diseases, and cancer.
XX
PS   Claim 25; Page 284; 0pp; English.
XX
CC   The invention relates to an isolated epitope present on cancer cells and
CC   important in physiological phenomena such as cell rolling, metastasis and
CC   inflammation, where the epitope is capable of being bound by an antibody,
CC   its antigen-binding fragment or its complex comprising at least one
CC   antibody or its binding fragment having a first hypervariable region. The
CC   epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
CC   disease, thrombosis, restenosis, metastasis, growth and/or replication of
CC   tumour or leukaemia cells, increase in number of tumour or leukaemia
CC   cells in a patient, cell-cell, cell-matrix, platelet-matrix, platelet-
CC   platelet and/or cell-platelet adhesion or aggregation, for increasing
CC   mortality of tumour or leukaemia cells, for increasing the susceptibility
CC   of diseased cells to damage by anti-disease, anti-cancer or anti-
CC   leukaemia agents, or for decreasing the number of tumour or leukaemia
CC   cells in a patient, or in the manufacture of a medicament for the above
CC   mentioned purposes. The epitopes are useful for diagnosing and treating
CC   diseases such as cancer, leukaemia, autoimmune diseases, inflammatory
CC   diseases, cardiovascular diseases such as myocardial infarction,
CC   retinopathic diseases and other diseases mediated by abnormal platelet
CC   function and diseases caused by sulphated tyrosine-dependent protein-
CC   protein interactions. This sequence represents a human antibody fragment
CC   of the invention
XX
```

SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
||||||
Db 1 DYGMS 5

RESULT 4
ABU11271
ID ABU11271 standard; peptide; 5 AA.
XX
AC ABU11271;
XX
DT 06-FEB-2003 (first entry)
XX
DE Human TANGO 268 VHCDR1 Peptide #5.
XX
KW Human; mouse; variable heavy; VH; antigen; cancer;
KW complementarity determining region; TANGO 268; glycoprotein VI; GPVI;
KW TANGO 268; extracellular matrix; collagen; platelet release;
KW proliferation; migration; embryogenesis; inflammation; thrombosis;
KW degranulation; thrombocytopaenia; antibody; thrombotic disorder;
KW cerebral vascular disease; stroke; ischaemia; venous thromboembolism;
KW leg swelling; pain; ulceration; pulmonary embolism; coronary disease;
KW cardiovascular disease; angina pectoris; myocardial infarction;
KW coronary restenosis; atherosclerosis; immunological disorder;
KW developmental disorder; embryonic disorder; liver disorder;
KW cerebral vascular disease; venous thromboembolism disease.
XX
OS Homo sapiens.
XX
PN WO200280968-A1.
XX
PD 17-OCT-2002.
XX
PF 09-APR-2002; 2002WO-US011122.
XX
PR 09-APR-2001; 2001US-00829495.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Busfield SJ, Villeval J, Jandrot-Perrus M, Vainchencker W;
PI Gill DS, Qian DM, Kingsbury G;
XX
DR WPI; 2003-058477/05.
XX
PT Novel substantially purified antibody immunospecifically binding to TANGO
PT 268 antigen, useful for treating bleeding disorders such as
PT thrombocytopenia, stroke, ischemia, pulmonary embolism, atherosclerosis.
XX
PS Claim 1; Page 111; 236pp; English.
XX
CC This invention relates to a novel purified antibody comprising a variable
CC heavy (VH) complementarity determining region (CDR)1, VH CDR2 or VH CDR3;
CC or variable light (VL) CDR1, VL CDR2 or VL CDR3, and immunospecifically
CC binding to a TANGO 268 (also referred as glycoprotein VI (GPVI)) antigen.
CC The antibodies of the invention act to decrease or block TANGO 268
CC binding to extracellular matrix components, or as a Collagen or platelet
CC release and aggregation blocker. The antibodies of the invention are
CC useful for modulating proliferation, migration, morphology,
CC differentiation and/or function of megakaryocytes and platelets,
CC including during development e.g. embryogenesis, modulating leukocyte-
CC platelet and platelet-endothelium interactions in inflammation and/or
CC thrombosis, and modulating platelet aggregation and degranulation. They
CC are also useful for modulating disorders associated with abnormal or
CC aberrant megakaryocyte and/or platelet proliferation, migration,
CC morphology, differentiation and/or function, e.g. bleeding disorders such
CC as thrombocytopaenia. Other diseases which may be modulated by these
CC antibodies are thrombotic disorders, cerebral vascular diseases (e.g.
CC stroke and ischaemia) venous thromboembolism diseases (e.g. diseases
CC involving leg swelling, pain and ulceration, pulmonary embolism, etc);
CC coronary diseases (e.g. cardiovascular diseases including angina
CC pectoris, myocardial infarction, coronary restenosis, atherosclerosis,
CC etc); immunological disorders, developmental disorders, embryonic
CC disorders, liver disorders, cerebral vascular diseases, venous
CC thromboembolism disease, coronary diseases, and metastatic cancers. The
CC antibodies of the invention only causes a transient decrease in platelet
CC counts, platelet aggregation, and/or platelet activation and so have some
CC advantages over prior art methods. The present sequence represents a
CC peptide sequence used to generate the antibodies of the invention
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 6; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
||||||
Db 1 DYGMS 5

RESULT 5
AAR49460
ID AAR49460 standard; protein; 16 AA.
XX
AC AAR49460;
XX
DT 25-MAR-2003 (revised)
DT 16-SEP-1994 (first entry)
XX
DE Factor VIII position 1175-1790.
XX
KW Naturally-occuring; immunomodulatory protein; human; therapy; class I;
KW major histocompatibility complex; class II; allotype; type I diabetes;
KW autoimmune disease; rheumatoid arthritis; T-cell-mediated response;
KW mutiple sclerosis; transplant rejection; vaccine; MHC.
XX
OS Homo sapiens.
XX
PN WO9404171-A1.
XX
PD 03-MAR-1994.
XX
PF 11-AUG-1993; 93WO-US007545.
XX
PR 11-AUG-1992; 92US-00925460.
PR 15-JUN-1993; 93US-00077255.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Urban RG, Chicz RM, Vignali DA, Hedley ML, Stern LJ;
PI Strominger JL;
XX
DR WPI; 1994-082825/10.
XX
PT Novel immunomodulatory peptide(s) and nucleic acids - useful for
PT treatment of auto:immune diseases, transplant rejection and for
PT vaccination.
XX
PS Disclosure; Page 41; 139pp; English.
XX
CC The sequences given in AAR49291-505 and AAR46981-7038 represent peptide
CC fragments of naturally-occuring immunomodulatory proteins. These
CC fragments are between 10-30 residues in length and bind to a human major
CC histocompatibility complex (MHC) class II allotype. These peptides may be
CC used for therapy of autoimmune diseases, such as type I diabetes,
CC rheumatoid arthritis and mutiple sclerosis, and to reduce transplant
CC rejection. They may also be used for vaccination providing an exclusively
CC T-cell-mediated response, which can be class I or class-II based, or
CC both, depending on the length and character of the immunogenic peptides.
CC (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to

```
CC   correct PR field.)
XX
SQ   Sequence 16 AA;

  Query Match             100.0%;  Score 28;  DB 2;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          3 DYGMS 7


RESULT 6
ABJ05002
ID   ABJ05002 standard; peptide; 20 AA.
XX
AC   ABJ05002;
XX
DT   30-OCT-2002  (first entry)
XX
DE   A3 peptide of fVII precursor SEQ ID No 58.
XX
KW   Haemostatic; antibody inhibitor; factor VIII; T cell; immune response;
KW   haemophilia A; acquired haemophilia; human factor VIII.
XX
OS   Homo sapiens.
XX
PN   WO200260917-A2.
XX
PD   08-AUG-2002.
XX
PF   30-NOV-2001; 2001WO-US044945.
XX
PR   01-DEC-2000; 2000US-0250430P.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Conti-Fine BM;
XX
DR   WPI; 2002-627462/67.
XX
PT   New peptides or their variants, useful for preventing, treating or
PT   inhibiting aberrant or pathogenic production of antibodies specific for
PT   factor VIII, particularly useful for treating hemophilia A or acquired
PT   hemophilia.
XX
PS   Claim 44; Page 62; 120pp; English.
XX
CC   The invention relates to isolated and purified peptides and variants
CC   thereof, as well as DNA encoding those peptides, useful to prevent or
CC   treat antibody inhibitors of factor VIII. The peptides are useful for
CC   preventing or inhibiting aberrant, pathogenic or undesirable antibody
CC   production or antibody binding that is specific for factor VIII. The
CC   peptides are also useful for preventing or inhibiting the priming or
CC   activity of T cells specific for factor VIII. These peptides are also
CC   useful for enhancing the activity or increasing the levels of modulatory
CC   T cells that inhibit the immune response to factor VIII. These peptides
CC   are useful in mammals, particularly in humans. The administration of
CC   these peptides does not increase the synthesis of a pathogenic antibody
CC   to factor VIII, or its biologically active fragment or functional
CC   equivalent. In particular, these peptides are useful for treating
CC   haemophilia A or acquired haemophilia. This sequence represents a human
CC   factor VIII peptide region relating to the invention
XX
SQ   Sequence 20 AA;

  Query Match             100.0%;  Score 28;  DB 5;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          9 DYGMS 13


RESULT 7
AAP61435
ID   AAP61435 standard; protein; 26 AA.
XX
AC   AAP61435;
XX
DT   25-MAR-2003  (revised)
DT   02-AUG-1991  (first entry)
XX
DE   Factor VIII:C coagulant polypeptide.
XX
KW   Factor VIII:C; coagulant; haemophilia; clotting; blood.
XX
OS   Homo sapiens.
XX
PN   EP182372-A.
XX
PD   28-MAY-1986.
XX
PF   19-NOV-1985;   85EP-00114743.
XX
PR   31-MAR-1983;   83US-00481105.
PR   30-NOV-1983;   83US-00556508.
PR   21-NOV-1984;   84US-00673916.
PR   24-MAY-1985;   85US-00738134.
PR   13-APR-1987;   87US-00037615.
XX
PA   (SCRI ) SCRIPPS CLINIC & RES FOUND.
XX
PI   Zimmermann TS,  Fulcher CA;
XX
DR   WPI; 1986-138803/22.
XX
PT   New factor VIII:c coagulant polypeptide(s) - useful for enhanced activity
PT   in treating haemophilia and obtd. by treating factor VIII:c with
PT   protease.
XX
PS   Claim 6; Page 44; 46pp; English.
XX
CC   The polypeptide has mol. wt. of approx. 79000-80000. It is an active
CC   Factor VIII:C coagulant polypeptide, and has enhanced coagulant activity.
CC   It is useful for treating haemophilia. It is also useful in the study and
CC   characterisation of polypeptides or their complexes providing desired
CC   clotting behaviour in the blood of humans and other mammals. Activity may
CC   be 3-5 times that of known purified human Factor VIII:C, or even 10-100
CC   times. Activity is shown over a continuous period of at least 10 mins.,
CC   and usually for much longer. The polypeptide may be mixed with other
CC   FActor VIII:C polypeptides to produce a specific coagulant activity of
CC   over 1000 units/ml. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 26 AA;

  Query Match             100.0%;  Score 28;  DB 1;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         19 DYGMS 23


RESULT 8
AAP50314
ID   AAP50314 standard; protein; 29 AA.
XX
AC   AAP50314;
XX
DT   25-MAR-2003  (revised)
DT   23-NOV-1991  (first entry)
XX
```

DE Peptide encoded by porcine antihaemophilic factor exon (34-H1).
XX
KW Porcine antihaemophilic factor; haemophilia; factor VIIIC.
XX
OS Sus scrofa.
XX
PN WO8501961-A.
XX
PD 09-MAY-1985.
XX
PF 12-OCT-1984; 84WO-US001641.
XX
PR 28-OCT-1983; 83US-00546650.
PR 24-AUG-1984; 84US-00644036.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Toole J;
XX
DR WPI; 1985-122479/20.
XX
PT Prodn. of human and porcine factor VIIIc - by using recombinant dna
PT techniques for their cellular prodn.
XX
PS Disclosure; Fig 1B; 71pp; English.
XX
CC The peptide is encoded by an exon of the porcine antihaemophilic factor
CC (AHF) gene. The exon, from nucleotides 169 to 267, encodes at least 30
CC amino acids from Phe 2 ro Arg 31 of the 69 kD fragment. The AHF was obtd.
CC using anti-factor VIIIC monoclonal antibody. The peptide is useful for
CC the treatment of haemophilia. (Updated on 25-MAR-2003 to correct PR
CC field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 29 AA;

Query Match 100.0%; Score 28; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 20 DYGMS 24

RESULT 9
AAP50318
ID AAP50318 standard; protein; 33 AA.
XX
AC AAP50318;
XX
DT 25-MAR-2003 (revised)
DT 23-NOV-1991 (first entry)
XX
DE Peptide encoded by HaeIII insert 34-H1 bearing porcine antihaemophilic
DE factor exon.
XX
KW Porcine antihaemophilic factor; haemophilia; insert 34-H1; factor VIIIC.
XX
OS Sus scrofa.
XX
FH Key Location/Qualifiers
FT Region 3. .33
FT /label= human factor VIIIC sequence
XX
PN WO8501961-A.
XX
PD 09-MAY-1985.
XX
PF 12-OCT-1984; 84WO-US001641.
XX
PR 28-OCT-1983; 83US-00546650.
PR 24-AUG-1984; 84US-00644036.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Toole J;
XX
DR WPI; 1985-122479/20.
DR N-PSDB; AAN50373.
XX
PT Prodn. of human and porcine factor VIIIc - by using recombinant dna
PT techniques for their cellular prodn.
XX
PS Disclosure; Fig 5; 71pp; English.
XX
CC The peptide is encoded by the HaeIII insert 34-H1, which bears the exon
CC for porcine antihaemophilic factor (AHF). It contains a portion of human
CC factor VIIIC. The nucleotide sequence encoding this human region
CC (claimed) can be used as a screening agent to identify DNA and RNA
CC sequences encoding at least part of the human factor VIIIC gene. Human
CC factor VIIIC and porcine AHF are useful for the treatment of haemophilia.
CC (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to
CC correct PA field.)
XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 28; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 20 DYGMS 24

RESULT 10
AAP50313
ID AAP50313 standard; peptide; 41 AA.
XX
AC AAP50313;
XX
DT 25-MAR-2003 (revised)
DT 23-NOV-1991 (first entry)
XX
DE Thrombin cleavage prod. of porcine antihaemophilic factor.
XX
KW Porcine antihaemophilic factor; haemophilia; thrombin cleavage prod;
KW factor VIIIC.
XX
OS Sus scrofa.
XX
PN WO8501961-A.
XX
PD 09-MAY-1985.
XX
PF 12-OCT-1984; 84WO-US001641.
XX
PR 28-OCT-1983; 83US-00546650.
PR 24-AUG-1984; 84US-00644036.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Toole J;
XX
DR WPI; 1985-122479/20.
XX
PT Prodn. of human and porcine factor VIIIc - by using recombinant dna
PT techniques for their cellular prodn.
XX
PS Disclosure; Fig 1A; 71pp; English.
XX
CC The peptide is the N-terminal sequence of the 69 kD thrombin cleavage
CC prod. of porcine antihaemophilic factor (AHF). The AHF was obtd. using
CC anti-factor VIIIC monoclonal antibody. The peptide is useful for the
CC treatment of haemophilia. (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PA field.)

```
XX
SQ   Sequence 41 AA;

  Query Match             100.0%;  Score 28;  DB 1;  Length 41;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         18 DYGMS 22


RESULT 11
ABJ04948
ID   ABJ04948 standard; protein; 60 AA.
XX
AC   ABJ04948;
XX
DT   30-OCT-2002  (first entry)
XX
DE   A3 domain of fVII epitope SEQ ID No 4.
XX
KW   Haemostatic; antibody inhibitor; factor VIII; T cell; immune response;
KW   haemophilia A; acquired haemophilia; human factor VIII.
XX
OS   Homo sapiens.
XX
PN   WO200260917-A2.
XX
PD   08-AUG-2002.
XX
PF   30-NOV-2001; 2001WO-US044945.
XX
PR   01-DEC-2000; 2000US-0250430P.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Conti-Fine BM;
XX
DR   WPI; 2002-627462/67.
XX
PT   New peptides or their variants, useful for preventing, treating or
PT   inhibiting aberrant or pathogenic production of antibodies specific for
PT   factor VIII, particularly useful for treating hemophilia A or acquired
PT   hemophilia.
XX
PS   Claim 4; Page 84; 120pp; English.
XX
CC   The invention relates to isolated and purified peptides and variants
CC   thereof, as well as DNA encoding those peptides, useful to prevent or
CC   treat antibody inhibitors of factor VIII. The peptides are useful for
CC   preventing or inhibiting aberrant, pathogenic or undesirable antibody
CC   production or antibody binding that is specific for factor VIII. The
CC   peptides are also useful for preventing or inhibiting the priming or
CC   activity of T cells specific for factor VIII. These peptides are also
CC   useful for enhancing the activity or increasing the levels of modulatory
CC   T cells that inhibit the immune response to factor VIII. These peptides
CC   are useful in mammals, particularly in humans. The administration of
CC   these peptides does not increase the synthesis of a pathogenic antibody
CC   to factor VIII, or its biologically active fragment or functional
CC   equivalent. In particular, these peptides are useful for treating
CC   haemophilia A or acquired haemophilia. This sequence represents a human
CC   factor VIII protein region relating to the invention
XX
SQ   Sequence 60 AA;

  Query Match             100.0%;  Score 28;  DB 5;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         39 DYGMS 43


RESULT 12
AAB07206
ID   AAB07206 standard; protein; 73 AA.
XX
AC   AAB07206;
XX
DT   11-OCT-2000  (first entry)
XX
DE   Human Factor VIIIC 77/80kd subunit partial protein sequence #1.
XX
KW   Factor VIII C domain; human; blood clotting; haemophilia.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 53
FT                   /note= "encoded by CGT"
XX
PN   EP1006182-A2.
XX
PD   07-JUN-2000.
XX
PF   11-JAN-1985; 2000EP-00200860.
XX
PR   12-JAN-1984;   84US-00570062.
PR   26-OCT-1984;   84US-00664919.
PR   11-JAN-1985;   85EP-00100223.
PR   11-JAN-1985;   91EP-00113267.
XX
PA   (CHIR ) CHIRON CORP.
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Kuo G,  Rasmussen ME,  Masiarz FR,  Valenzuela P,  Truett M;
PI   Favaloro J;
XX
DR   WPI; 2000-367967/32.
DR   N-PSDB; AAA58444.
XX
PT   DNA sequence of portion of human Factor VIIIC for treating and preventing
PT   the symptoms of hemophilia.
XX
PS   Claim 3; Page 37; 39pp; English.
XX
CC   The present sequence is the protein sequence for the 77/80kd subunit of
CC   human Factor VIIIC. Factor VIIIC is a plasma protein involved in blood
CC   coagulation, and is absent or defective in haemophilia A. The Factor
CC   VIIIC protein can, therefore, be used to treat haemophilia, as well as in
CC   the production of monoclonal antibodies to Factor VIIIC, and in
CC   diagnostic assays for the presence of Factor VIIIC subunits in
CC   physiological fluids, for example blood or serum
XX
SQ   Sequence 73 AA;

  Query Match             100.0%;  Score 28;  DB 3;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 62;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         60 DYGMS 64


RESULT 13
AAB40073
ID   AAB40073 standard; protein; 98 AA.
XX
AC   AAB40073;
XX
DT   05-FEB-2001  (first entry)
XX
DE   Anti-hIL12 antibody H chain V region amino acid sequence SEQ ID 599.
```

XX
KW Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
KW complementarity determining region; CDR; antirheumatic; antiarthritic;
KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiant;
KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200056772-A1.
XX
PD 28-SEP-2000.
XX
PF 24-MAR-2000; 2000WO-US007946.
XX
PR 25-MAR-1999; 99US-0126603P.
XX
PA (BADI) BASF AG.
PA (GEMY) GENETICS INST INC.
XX
PI Salfeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
PI Kaymakcalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
PI Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR;
PI Derbyshire EJ, Carmen S, Smith S, Holtet TL, Du Fou SL;
XX
DR WPI; 2000-638250/61.
XX
PT New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT disease and multiple sclerosis.
XX
PS Claim 75; Page 121; 377pp; English.
XX
CC This invention relates to a new human antibody specific for human
CC interleukin-12 (IL-12). The invention also includes antigen binding
CC portions that bind to IL-12. Sequences AAB39485-B39516 represent human
CC anti-IL-12 antibody heavy and light chain complementarity determining
CC region (CDR) amino acid sequences, and also includes variable region
CC amino acid sequences. Other variable region amino acid sequences are
CC given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771
CC represent anti-IL-12 CDR3 related amino acid sequences, AAB39772-B40063
CC represent other CDR sequences. Light chain CDR3 consensus sequences are
CC given in AAB40064-B40067. Primers used in the identification and
CC construction of the antibodies of the invention are given in AAC61062-
CC C61071. The antibody of the invention is a neutralising antibody and has
CC antirheumatic; antiarthritic; antisclerotic; antiinflammatory;
CC neuroprotective; antipsoriatic; antiasthmatic; cardiant; antiparasitic;
CC antibacterial and immunosuppressive activity. The antibodies or antigen-
CC binding fragments are useful in the treatment of disorders associated
CC with detrimental release of human IL-12, especially Crohn's disease,
CC multiple sclerosis and rheumatoid arthritis. They can also be used in the
CC manufacture of a pharmaceutical composition to treat human IL-12
CC disorders
XX
SQ Sequence 98 AA;

Query Match 100.0%; Score 28; DB 3; Length 98;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 31 DYGMS 35

RESULT 14
ABG78186
ID ABG78186 standard; protein; 98 AA.
XX
AC ABG78186;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human Fv molecule hypervariable region related peptide #61.
XX
KW Human; Fv molecule; hypervariable region; single chain Fv; cytostatic;
KW disulfide Fv; dsFv; scFv; cancer; carcinoma; sarcoma; leukaemia; adenoma;
KW lymphoma; myeloma; blastoma; seminoma; melanoma; acute myeloid leukaemia.
XX
OS Homo sapiens.
XX
PN WO200259264-A2.
XX
PD 01-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049440.
XX
PR 29-DEC-2000; 2000US-00751181.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Hagai Y, Lazarovits J, Guy R, Lipschitz O, Szanton E, Levanon A;
PI Plaksin D, Peretz T;
XX
DR WPI; 2002-619166/66.
XX
PT Novel peptide/polypeptide for cancer therapy has Fv molecule, construct
PT or fragment, or construct of fragment with enhanced binding
PT characteristics so as to selectively bind target cell in favor of other
PT cells.
XX
PS Claim 13; Page 177-178; 232pp; English.
XX
CC The invention relates to a peptide or polypeptide comprising an Fv
CC molecule, a construct or fragments or a construct of a fragment with
CC enhanced binding characteristics which selectively and/or specifically
CC binds to a target cell in favour of other cells, where binding is
CC primarily determined by a first hypervariable region and Fv is a single
CC chain Fv (scFv) or a disulfide Fv (dsFv). The peptide, optionally in
CC association with or attached, coupled, combined, linked or fused to a
CC pharmaceutical agent, is useful in the manufacture of a medicament, where
CC the medicament has activity against a diseased cell, preferably a cancer
CC cell (selected from carcinoma, sarcoma, leukaemia, adenoma, lymphoma,
CC myeloma, blastoma, seminoma, and melanoma, where the leukaemia cell is an
CC acute myeloid leukaemia cell). The peptide is also useful for preparing a
CC composition for use in inhibiting the growth of a diseased or cancer
CC cell. This sequence represents a human Fv molecule hypervariable region
CC related peptide of the invention
XX
SQ Sequence 98 AA;

Query Match 100.0%; Score 28; DB 5; Length 98;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 31 DYGMS 35

RESULT 15
ABG91877
ID ABG91877 standard; protein; 98 AA.
XX
AC ABG91877;
XX
DT 04-DEC-2002 (first entry)
XX
DE Human antibody fragment #61.
XX
KW Human; antibody; epitope; cancer; tumour; cell rolling; inflammation;
KW metastasis; hypervariable region; autoimmune disease; thrombosis;
KW restenosis; leukaemia; inflammatory disease; cardiovascular disease;
KW myocardial infarction; retinopathic disease; abnormal platelet function;
KW sulphated tyrosine-dependent protein-protein interaction.
XX

OS Homo sapiens.
XX
PN WO200253700-A2.
XX
PD 11-JUL-2002.
XX
PF 31-DEC-2001; 2001WO-US049442.
XX
PR 29-DEC-2000; 2000US-00751181.
PR 29-DEC-2000; 2000US-0258948P.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Lazarovits J, Hagai Y, Plaksin D, Vogel T, Nimrod A, Mar-Haim H;
PI Szanthon E, Richter T, Amit B, Kooperman L, Peretz T, Levanon A;
XX
DR WPI; 2002-674776/72.
XX
PT Novel isolated epitope present on cancer cells and important in
PT physiological phenomena such as cell rolling, metastasis and
PT inflammation, for treating autoimmune, inflammatory or cardiovascular
PT diseases, and cancer.
XX
PS Disclosure; Page 255; 0pp; English.
XX
CC The invention relates to an isolated epitope present on cancer cells and
CC important in physiological phenomena such as cell rolling, metastasis and
CC inflammation, where the epitope is capable of being bound by an antibody,
CC its antigen-binding fragment or its complex comprising at least one
CC antibody or its binding fragment having a first hypervariable region. The
CC epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
CC disease, thrombosis, restenosis, metastasis, growth and/or replication of
CC tumour or leukaemia cells, increase in number of tumour or leukaemia
CC cells in a patient, cell-cell, cell-matrix, platelet-matrix, platelet-
CC platelet and/or cell-platelet adhesion or aggregation, for increasing
CC mortality of tumour or leukaemia cells, for increasing the susceptibility
CC of diseased cells to damage by anti-disease, anti-cancer or anti-
CC leukaemia agents, or for decreasing the number of tumour or leukaemia
CC cells in a patient, or in the manufacture of a medicament for the above
CC mentioned purposes. The epitopes are useful for diagnosing and treating
CC diseases such as cancer, leukaemia, autoimmune diseases, inflammatory
CC diseases, cardiovascular diseases such as myocardial infarction,
CC retinopathic diseases and other diseases mediated by abnormal platelet
CC function and diseases caused by sulphated tyrosine-dependent protein-
CC protein interactions. This sequence represents a human antibody fragment
CC of the invention
XX
SQ Sequence 98 AA;

Query Match 100.0%; Score 28; DB 5; Length 98;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 DYGMS 5
              |||||
Db         31 DYGMS 35
```

Search completed: September 18, 2004, 03:42:53
Job time : 37.3571 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 03:39:30 ; Search time 11.6071 Seconds
(without alignments)
22.239 Million cell updates/sec

Title: US-10-029-926B-114
Perfect score: 28
Sequence: 1 DYGMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	100.0	16	2	US-08-480-190-170	Sequence 170, App
2	28	100.0	16	2	US-08-488-379-170	Sequence 170, App
3	28	100.0	16	4	US-08-475-399A-170	Sequence 170, App
4	28	100.0	16	5	PCT-US93-07545-170	Sequence 170, App
5	28	100.0	116	2	US-08-652-816A-14	Sequence 14, Appl
6	28	100.0	117	3	US-08-545-809A-106	Sequence 106, App
7	28	100.0	232	4	US-09-489-039A-8586	Sequence 8586, Ap
8	28	100.0	258	4	US-09-107-532A-4002	Sequence 4002, Ap
9	28	100.0	264	4	US-09-107-532A-6813	Sequence 6813, Ap
10	28	100.0	309	4	US-09-079-029-9	Sequence 9, Appli
11	28	100.0	354	4	US-08-724-984A-2	Sequence 2, Appli
12	28	100.0	474	4	US-09-486-192-4	Sequence 4, Appli
13	28	100.0	868	1	US-07-864-004B-6	Sequence 6, Appli
14	28	100.0	868	1	US-08-251-937A-6	Sequence 6, Appli
15	28	100.0	868	1	US-08-212-133A-3	Sequence 3, Appli
16	28	100.0	1090	5	PCT-US93-03275-6	Sequence 6, Appli
17	28	100.0	1438	4	US-09-209-916-1	Sequence 1, Appli
18	28	100.0	1443	2	US-08-670-707A-39	Sequence 39, Appl
19	28	100.0	1443	3	US-09-037-601-39	Sequence 39, Appl
20	28	100.0	1443	4	US-09-315-179-39	Sequence 39, Appl
21	28	100.0	1467	4	US-09-523-656-38	Sequence 38, Appl
22	28	100.0	1471	1	US-08-683-839B-3	Sequence 3, Appli
23	28	100.0	1661	2	US-08-882-083-2	Sequence 2, Appli
24	28	100.0	1661	2	US-08-558-107-2	Sequence 2, Appli
25	28	100.0	1661	3	US-09-243-539-2	Sequence 2, Appli
26	28	100.0	2115	3	US-09-324-867-5	Sequence 5, Appli
27	28	100.0	2133	2	US-08-670-707A-37	Sequence 37, Appl
28	28	100.0	2133	3	US-09-037-601-37	Sequence 37, Appl
29	28	100.0	2133	4	US-09-315-179-37	Sequence 37, Appl
30	28	100.0	2133	4	US-09-523-656-30	Sequence 30, Appl
31	28	100.0	2304	3	US-09-324-867-4	Sequence 4, Appli
32	28	100.0	2319	1	US-08-212-133A-8	Sequence 8, Appli
33	28	100.0	2319	1	US-08-474-503-6	Sequence 6, Appli
34	28	100.0	2319	2	US-08-670-707A-6	Sequence 6, Appli
35	28	100.0	2319	3	US-09-037-601-6	Sequence 6, Appli
36	28	100.0	2319	4	US-09-315-179-6	Sequence 6, Appli
37	28	100.0	2319	4	US-09-523-656-28	Sequence 28, Appl
38	28	100.0	2319	5	PCT-US94-13200-6	Sequence 6, Appli
39	28	100.0	2332	1	US-07-864-004B-4	Sequence 4, Appli
40	28	100.0	2332	1	US-08-251-937A-4	Sequence 4, Appli
41	28	100.0	2332	1	US-08-212-133A-2	Sequence 2, Appli
42	28	100.0	2332	1	US-08-276-594A-2	Sequence 2, Appli
43	28	100.0	2332	1	US-08-474-503-2	Sequence 2, Appli
44	28	100.0	2332	2	US-08-670-707A-2	Sequence 2, Appli
45	28	100.0	2332	3	US-09-037-601-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-480-190-170
; Sequence 170, Application US/08480190
; Patent No. 5827516
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,190
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  June 15, 1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  August 11, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00246/168001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  170:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-480-190-170
```

```
  Query Match             100.0%;  Score 28;  DB 2;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 3;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          3 DYGMS 7


RESULT 2
US-08-488-379-170
; Sequence 170, Application US/08488379
; Patent No. 5880103
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,379
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  June 15, 1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  August 11, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00246/168001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  170:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-488-379-170

  Query Match             100.0%;  Score 28;  DB 2;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 3;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          3 DYGMS 7


RESULT 3
US-08-475-399A-170
; Sequence 170, Application US/08475399A
; Patent No. 6509033
;  GENERAL INFORMATION:
;    APPLICANT:  Urban, Robert G.
;    APPLICANT:  Chicz, Roman M.
;    APPLICANT:  Vignali, Dario A.A.
;    APPLICANT:  Hedley, Mary L.
;    APPLICANT:  Stern, Lawrence J.
;    APPLICANT:  Strominger, Jack L.
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  276
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,399A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  15-JUN-1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  11-AUG-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  00246/168003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-507
;      TELEFAX:  617/542-890
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  170:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-475-399A-170

  Query Match             100.0%;  Score 28;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 3;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          3 DYGMS 7


RESULT 4
PCT-US93-07545-170
; Sequence 170, Application PC/TUS9307545
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  273
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
```

; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07545
; FILING DATE: 19930811
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US93-07545-170

Query Match 100.0%; Score 28; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 3 DYGMS 7

RESULT 5
US-08-652-816A-14
; Sequence 14, Application US/08652816A
; Patent No. 5872215
; GENERAL INFORMATION:
; APPLICANT: Osbourn, JK
; APPLICANT: Allen, DJ
; APPLICANT: McCafferty, JG
; TITLE OF INVENTION: Specific binding members, materials and
; TITLE OF INVENTION: methods.
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,816A
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.4
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.8
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206318.9
; FILING DATE: 24-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206372.6
; FILING DATE: 23-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9525004.9
; FILING DATE: 07-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610824.6
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/02240
; FILING DATE: 02-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/244,597
; FILING DATE: 01-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: David W. Clough
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28111/33308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 116 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-652-816A-14

Query Match 100.0%; Score 28; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 31 DYGMS 35

RESULT 6
US-08-545-809A-106
; Sequence 106, Application US/08545809A
; Patent No. 6096878
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Matsuda, Fumihiko
; TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
; TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
; NUMBER OF SEQUENCES: 145
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,809A
; FILING DATE: 27-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP93/00603
; FILING DATE: 10-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 06501/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-545-809A-106

  Query Match             100.0%;  Score 28;  DB 3;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         50 DYGMS 54


RESULT 7
US-09-489-039A-8586
; Sequence 8586, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 8586
;   LENGTH: 232
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8586

  Query Match             100.0%;  Score 28;  DB 4;  Length 232;
  Best Local Similarity   100.0%;  Pred. No. 62;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        219 DYGMS 223


RESULT 8
US-09-107-532A-4002
; Sequence 4002, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
;             APPLICATION NUMBER: 60/051571
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (781)893-5007
;             TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 4002:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 258 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ORIGINAL SOURCE:
;             ORGANISM: Enterococcus faecium
;        FEATURE:
;             NAME/KEY:  misc_feature
;             LOCATION: (B) LOCATION 1...258
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4002:
US-09-107-532A-4002

  Query Match             100.0%;  Score 28;  DB 4;  Length 258;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        172 DYGMS 176


RESULT 9
US-09-107-532A-6813
; Sequence 6813, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
;             APPLICATION NUMBER: 60/051571
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (781)893-5007
;             TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 6813:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 264 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ORIGINAL SOURCE:
```

```
;           ORGANISM: Enterococcus faecium
;       FEATURE:
;           NAME/KEY:  misc_feature
;           LOCATION: (B) LOCATION 1...264
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6813:
US-09-107-532A-6813

  Query Match             100.0%;  Score 28;  DB 4;  Length 264;
  Best Local Similarity   100.0%;  Pred. No. 72;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        178 DYGMS 182


RESULT 10
US-09-079-029-9
; Sequence 9, Application US/09079029
; Patent No. 6342369
;  GENERAL INFORMATION:
;    APPLICANT:  Adams, Camilia W.
;    APPLICANT:  Ashkenazi, Avi J.
;    APPLICANT:  Chuntharapai, Anan
;    APPLICANT:  Kim, Kyung J.
;    TITLE OF INVENTION:  Apo-2 Receptor
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  1 DNA Way
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,029
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Marschang, Diane L.
;      REGISTRATION NUMBER:  35,600
;      REFERENCE/DOCKET NUMBER:  P1101R2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650/225-5416
;      TELEFAX:  650/952-9881
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  309 amino acids
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
US-09-079-029-9

  Query Match             100.0%;  Score 28;  DB 4;  Length 309;
  Best Local Similarity   100.0%;  Pred. No. 86;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         70 DYGMS 74


RESULT 11
US-08-724-984A-2
; Sequence 2, Application US/08724984A
; Patent No. 6388055
;  GENERAL INFORMATION:
;    APPLICANT:  Derk Bergsma, Mary Brawner, and Usman Shabon
;    TITLE OF INVENTION:  No. 6388055el Mouse Genomic Clone of the CC-
;    TITLE OF INVENTION:  CKR5 Receptor
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road, P.O. Box 1539
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM 486
;      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;      SOFTWARE:  MICROSOFT WORD
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,984A
;      FILING DATE:  October 3, 1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  William T. Han
;      REGISTRATION NUMBER:  34,344
;      REFERENCE/DOCKET NUMBER:  ATG50023
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610 270 5024
;      TELEFAX:  610 270 5090
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  354
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
US-08-724-984A-2

  Query Match             100.0%;  Score 28;  DB 4;  Length 354;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         15 DYGMS 19


RESULT 12
US-09-486-192-4
; Sequence 4, Application US/09486192
; Patent No. 6521440
; GENERAL INFORMATION:
;  APPLICANT:  Estell, David A.
;  TITLE OF INVENTION:  Proteases From Gram-Positive Organisms
;  FILE REFERENCE:  GC386-US
;  CURRENT APPLICATION NUMBER: US/09/486,192
;  CURRENT FILING DATE:  2000-02-22
;  PRIOR APPLICATION NUMBER:  PCT/US98/18677
;  PRIOR FILING DATE:  1998-09-08
;  PRIOR APPLICATION NUMBER:  EP9719637.2
;  PRIOR FILING DATE:  1997-09-15
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 474
;   TYPE: PRT
;   ORGANISM: Schizocaccharomyces pombe
US-09-486-192-4

  Query Match             100.0%;  Score 28;  DB 4;  Length 474;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        349 DYGMS 353
```

```
RESULT 13
US-07-864-004B-6
; Sequence 6, Application US/07864004B
; Patent No. 5364771
;  GENERAL INFORMATION:
;    APPLICANT:  Lollar, John S.
;    APPLICANT:  Runge, Marschall S.
;    TITLE OF INVENTION:  Hybrid Human/Porcine Factor VIII
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/864,004B
;      FILING DATE:  07 APRIL 1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  EMU106
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-815-6508
;      TELEFAX:  404-815-6555
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  868 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Porcine
US-07-864-004B-6

  Query Match             100.0%;  Score 28;  DB 1;  Length 868;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        804 DYGMS 808


RESULT 14
US-08-251-937A-6
; Sequence 6, Application US/08251937A
; Patent No. 5583209
;  GENERAL INFORMATION:
;    APPLICANT:  Lollar, John S.
;    APPLICANT:  Runge, Marschall S.
;    TITLE OF INVENTION:  Hybrid Human/Porcine Factor VIII
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/251,937A
;      FILING DATE:  31-MAY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/864,004
;      FILING DATE:  07-APR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pratt, John S.
;      REGISTRATION NUMBER:  29,476
;      REFERENCE/DOCKET NUMBER:  EMU106DIV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-815-6367
;      TELEFAX:  404-815-6555
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  868 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Porcine
US-08-251-937A-6

  Query Match             100.0%;  Score 28;  DB 1;  Length 868;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        804 DYGMS 808


RESULT 15
US-08-212-133A-3
; Sequence 3, Application US/08212133A
; Patent No. 5663060
;  GENERAL INFORMATION:
;    APPLICANT:  Lollar, John S.
;    APPLICANT:  Runge, Marschall S.
;    TITLE OF INVENTION:  Hybrid Human/Animal Factor VIII
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  100 Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,133A
;      FILING DATE:  March 11, 1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/864,004
;      FILING DATE:  07-APR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
```

```
;      REFERENCE/DOCKET NUMBER:  EMU/76677
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-572-6508
;      TELEFAX:  404-572-6555
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  868 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Porcine
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..868
;      OTHER INFORMATION:  /note= "Predicted amino acid
;      OTHER INFORMATION:  sequence of the B and part of the A2 domains of
;      OTHER INFORMATION:  porcine factor VIII."
US-08-212-133A-3

  Query Match             100.0%;  Score 28;  DB 1;  Length 868;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        804 DYGMS 808


Search completed: September 18, 2004, 03:48:46
Job time : 11.6071 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 03:46:39 ; Search time 198.571 Seconds
(without alignments)
8.086 Million cell updates/sec

Title: US-10-029-926B-114
Perfect score: 28
Sequence: 1 DYGMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	100.0	5	9	US-09-832-312-73	Sequence 73, Appl
2	28	100.0	5	11	US-09-829-495-73	Sequence 73, Appl
3	28	100.0	5	12	US-10-029-926B-114	Sequence 114, App
4	28	100.0	5	15	US-10-032-037B-114	Sequence 114, App
5	28	100.0	5	15	US-10-029-988B-114	Sequence 114, App
6	28	100.0	5	15	US-10-032-423A-114	Sequence 114, App
7	28	100.0	20	16	US-10-433-273-58	Sequence 58, Appl
8	28	100.0	60	16	US-10-433-273-4	Sequence 4, Appli
9	28	100.0	98	12	US-10-453-698-60	Sequence 60, Appl
10	28	100.0	98	12	US-10-029-926B-61	Sequence 61, Appl
11	28	100.0	98	14	US-10-194-975-20	Sequence 20, Appl
12	28	100.0	98	15	US-10-308-817-60	Sequence 60, Appl
13	28	100.0	98	15	US-10-032-037B-61	Sequence 61, Appl
14	28	100.0	98	15	US-10-029-988B-61	Sequence 61, Appl
15	28	100.0	98	15	US-10-032-423A-61	Sequence 61, Appl
16	28	100.0	98	16	US-10-379-392-21	Sequence 21, Appl
17	28	100.0	116	16	US-10-437-963-114282	Sequence 114282,
18	28	100.0	156	12	US-10-424-599-215325	Sequence 215325,
19	28	100.0	239	10	US-09-880-748-937	Sequence 937, App
20	28	100.0	239	10	US-09-880-748-2015	Sequence 2015, Ap
21	28	100.0	239	10	US-09-880-748-2038	Sequence 2038, Ap
22	28	100.0	239	12	US-10-293-418-937	Sequence 937, App
23	28	100.0	239	12	US-10-293-418-2015	Sequence 2015, Ap
24	28	100.0	239	12	US-10-293-418-2038	Sequence 2038, Ap
25	28	100.0	244	14	US-10-322-673-42	Sequence 42, Appl
26	28	100.0	251	10	US-09-880-748-1542	Sequence 1542, Ap
27	28	100.0	251	12	US-10-293-418-1542	Sequence 1542, Ap
28	28	100.0	266	15	US-10-032-037B-204	Sequence 204, App
29	28	100.0	266	15	US-10-029-988B-204	Sequence 204, App
30	28	100.0	266	15	US-10-032-423A-204	Sequence 204, App
31	28	100.0	276	12	US-10-425-114-70003	Sequence 70003, A
32	28	100.0	277	12	US-10-029-926B-25	Sequence 25, Appl
33	28	100.0	277	12	US-10-029-926B-203	Sequence 203, App
34	28	100.0	277	15	US-10-032-037B-25	Sequence 25, Appl
35	28	100.0	277	15	US-10-032-037B-203	Sequence 203, App
36	28	100.0	277	15	US-10-029-988B-25	Sequence 25, Appl
37	28	100.0	277	15	US-10-029-988B-203	Sequence 203, App
38	28	100.0	277	15	US-10-032-423A-25	Sequence 25, Appl
39	28	100.0	277	15	US-10-032-423A-203	Sequence 203, App
40	28	100.0	279	12	US-10-425-114-36951	Sequence 36951, A
41	28	100.0	299	16	US-10-767-701-44522	Sequence 44522, A
42	28	100.0	300	12	US-10-425-114-37904	Sequence 37904, A
43	28	100.0	309	13	US-10-052-798-9	Sequence 9, Appli
44	28	100.0	309	14	US-10-288-917-9	Sequence 9, Appli
45	28	100.0	309	15	US-10-423-448-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-832-312-73
; Sequence 73, Application US/09832312
; Patent No. US20010049829A1
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 73
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-73

Query Match 100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 1 DYGMS 5

RESULT 2
US-09-829-495-73
; Sequence 73, Application US/09829495

```
; Publication No. US20040001826A1
; GENERAL INFORMATION:
;  APPLICANT: Busfield SJ
;  APPLICANT:  Villeval J
;  APPLICANT:  Jandrot-Perrus M
;  APPLICANT:  Vainchenker W
;  APPLICANT:  Gill DS
;  APPLICANT:  Qian MD
;  TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
;  FILE REFERENCE: 7853-234
;  CURRENT APPLICATION NUMBER: US/09/829,495
;  CURRENT FILING DATE:  2001-04-09
;  PRIOR APPLICATION NUMBER: 09/610,118
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: 09/503,387
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: 09/454,824
;  PRIOR FILING DATE: 1999-12-06
;  PRIOR APPLICATION NUMBER: 09/345,468
;  PRIOR FILING DATE: 1999-06-30
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 73
;   LENGTH: 5
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-829-495-73

  Query Match             100.0%;  Score 28;  DB 11;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          1 DYGMS 5


RESULT 3
US-10-029-926B-114
; Sequence 114, Application US/10029926B
; Publication No. US20040073011A1
; GENERAL INFORMATION:
;  APPLICANT: HAGAY, et al.
;  TITLE OF INVENTION: SPECIFIC HUMAN ANTIBODIES FOR SELECTIVE CANCER THERAPY
;  FILE REFERENCE: 10793/50
;  CURRENT APPLICATION NUMBER: US/10/029,926B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 12/29/2000
;  NUMBER OF SEQ ID NOS: 203
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 114
;   LENGTH: 5
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-029-926B-114

  Query Match             100.0%;  Score 28;  DB 12;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          1 DYGMS 5


RESULT 4
US-10-032-037B-114
; Sequence 114, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/44
;  CURRENT APPLICATION NUMBER: US/10/032,037B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 114
;   LENGTH: 5
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-032-037B-114

  Query Match             100.0%;  Score 28;  DB 15;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          1 DYGMS 5


RESULT 5
US-10-029-988B-114
; Sequence 114, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/46
;  CURRENT APPLICATION NUMBER: US/10/029,988B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 114
;   LENGTH: 5
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-029-988B-114

  Query Match             100.0%;  Score 28;  DB 15;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          1 DYGMS 5


RESULT 6
US-10-032-423A-114
; Sequence 114, Application US/10032423A
; Publication No. US20040002450A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/45
;  CURRENT APPLICATION NUMBER: US/10/032,423A
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 12/29/2000
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 114
;   LENGTH: 5
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-032-423A-114
```

```
  Query Match             100.0%;  Score 28;  DB 15;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          1 DYGMS 5


RESULT 7
US-10-433-273-58
; Sequence 58, Application US/10433273
; Publication No. US20040096456A1
; GENERAL INFORMATION:
;  APPLICANT: Regents of the University of Minnesota
;  APPLICANT:  Conti-Fine, Bianca M.
;  TITLE OF INVENTION: Methods to Treat Hemophilia
;  FILE REFERENCE: 600.507WO1
;  CURRENT APPLICATION NUMBER: US/10/433,273
;  CURRENT FILING DATE:  2003-11-17
;  PRIOR APPLICATION NUMBER: US 60/250,430
;  PRIOR FILING DATE: 2000-12-01
;  NUMBER OF SEQ ID NOS: 61
; SEQ ID NO 58
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-433-273-58

  Query Match             100.0%;  Score 28;  DB 16;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db          9 DYGMS 13


RESULT 8
US-10-433-273-4
; Sequence 4, Application US/10433273
; Publication No. US20040096456A1
; GENERAL INFORMATION:
;  APPLICANT: Regents of the University of Minnesota
;  APPLICANT:  Conti-Fine, Bianca M.
;  TITLE OF INVENTION: Methods to Treat Hemophilia
;  FILE REFERENCE: 600.507WO1
;  CURRENT APPLICATION NUMBER: US/10/433,273
;  CURRENT FILING DATE:  2003-11-17
;  PRIOR APPLICATION NUMBER: US 60/250,430
;  PRIOR FILING DATE: 2000-12-01
;  NUMBER OF SEQ ID NOS: 61
; SEQ ID NO 4
;   LENGTH: 60
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-433-273-4

  Query Match             100.0%;  Score 28;  DB 16;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 77;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         39 DYGMS 43


RESULT 9
US-10-453-698-60
; Sequence 60, Application US/10453698
; Publication No. US20040038308A1
; GENERAL INFORMATION:
;  APPLICANT: Rother, Russell
;  TITLE OF INVENTION: HYBRID ANTIBODIES
;  FILE REFERENCE: 82 CIP (1087-37 CIP)
;  CURRENT APPLICATION NUMBER: US/10/453,698
;  CURRENT FILING DATE:  2003-06-03
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 60
;   LENGTH: 98
;   TYPE: PRT
;   ORGANISM: human
US-10-453-698-60

  Query Match             100.0%;  Score 28;  DB 12;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         31 DYGMS 35


RESULT 10
US-10-029-926B-61
; Sequence 61, Application US/10029926B
; Publication No. US20040073011A1
; GENERAL INFORMATION:
;  APPLICANT: HAGAY, et al.
;  TITLE OF INVENTION: SPECIFIC HUMAN ANTIBODIES FOR SELECTIVE CANCER THERAPY
;  FILE REFERENCE: 10793/50
;  CURRENT APPLICATION NUMBER: US/10/029,926B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 12/29/2000
;  NUMBER OF SEQ ID NOS: 203
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
;   LENGTH: 98
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-029-926B-61

  Query Match             100.0%;  Score 28;  DB 12;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         31 DYGMS 35


RESULT 11
US-10-194-975-20
; Sequence 20, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
;  APPLICANT: Foote, Jefferson
;  TITLE OF INVENTION: Super Humanized Antibodies
;  FILE REFERENCE: 501231.01
;  CURRENT APPLICATION NUMBER: US/10/194,975
;  CURRENT FILING DATE:  2002-10-10
;  PRIOR APPLICATION NUMBER: US 60/305,111
;  PRIOR FILING DATE: 2001-07-12
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
;   LENGTH: 98
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-194-975-20

  Query Match             100.0%;  Score 28;  DB 14;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
```

```
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         31 DYGMS 35


RESULT 12
US-10-308-817-60
; Sequence 60, Application US/10308817
; Publication No. US20030219861A1
; GENERAL INFORMATION:
;  APPLICANT: Rother, Russell
;  APPLICANT:  Wu, Dayang
;  TITLE OF INVENTION: HYBRID ANTIBODIES
;  FILE REFERENCE: 1087-37
;  CURRENT APPLICATION NUMBER: US/10/308,817
;  CURRENT FILING DATE:  2002-12-03
;  NUMBER OF SEQ ID NOS: 195
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
;   LENGTH: 98
;   TYPE: PRT
;   ORGANISM: human
US-10-308-817-60

  Query Match             100.0%;  Score 28;  DB 15;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         31 DYGMS 35


RESULT 13
US-10-032-037B-61
; Sequence 61, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/44
;  CURRENT APPLICATION NUMBER: US/10/032,037B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
;   LENGTH: 98
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-032-037B-61

  Query Match             100.0%;  Score 28;  DB 15;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         31 DYGMS 35


RESULT 14
US-10-029-988B-61
; Sequence 61, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/46
;  CURRENT APPLICATION NUMBER: US/10/029,988B
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
;   LENGTH: 98
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-029-988B-61

  Query Match             100.0%;  Score 28;  DB 15;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         31 DYGMS 35


RESULT 15
US-10-032-423A-61
; Sequence 61, Application US/10032423A
; Publication No. US20040002450A1
; GENERAL INFORMATION:
;  APPLICANT: Bio-Technology General Corp.
;  TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
;  TITLE OF INVENTION:  MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
;  FILE REFERENCE: 10793/45
;  CURRENT APPLICATION NUMBER: US/10/032,423A
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: 60/258,948
;  PRIOR FILING DATE: 12/29/2000
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
;   LENGTH: 98
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-032-423A-61

  Query Match             100.0%;  Score 28;  DB 15;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         31 DYGMS 35


Search completed: September 18, 2004, 04:20:42
Job time : 199.571 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 03:33:04 ; Search time 10.1786 Seconds
(without alignments)
47.252 Million cell updates/sec

Title: US-10-029-926B-114
Perfect score: 28
Sequence: 1 DYGMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	100.0	98	2	S26928	Ig heavy chain V r
2	28	100.0	112	2	PH1654	Ig heavy chain V r
3	28	100.0	119	2	A43413	Ig heavy chain V r
4	28	100.0	227	2	AB1076	probable fimbrial
5	28	100.0	268	2	H84684	En/Spm-like transp
6	28	100.0	268	2	T04660	hypothetical prote
7	28	100.0	334	2	H95307	arginine deiminase
8	28	100.0	340	2	H72681	hypothetical prote
9	28	100.0	411	2	S74760	hypothetical prote
10	28	100.0	448	2	AB3043	oxidoreductase Atu
11	28	100.0	468	2	B98243	hypothetical prote
12	28	100.0	474	2	T38737	probable CAAX pren
13	28	100.0	478	2	F89651	protein T04F8.2 [i
14	28	100.0	502	2	T24471	hypothetical prote
15	28	100.0	609	2	H71285	probable cell divi
16	28	100.0	650	2	G87572	calcium-binding pr
17	28	100.0	756	2	D96527	protein F27J15.24
18	28	100.0	830	2	I50455	prolactin receptor
19	28	100.0	869	2	A25945	coagulation factor
20	28	100.0	916	2	B84473	copia-like retroel
21	28	100.0	1008	2	H85055	probable transposo
22	28	100.0	1250	2	T27706	hypothetical prote
23	28	100.0	2133	2	T42763	coagulation factor
24	28	100.0	2319	2	A47004	coagulation factor
25	28	100.0	2351	1	EZHU	coagulation factor
26	25	89.3	109	2	T47696	hypothetical prote
27	25	89.3	113	2	B36259	Ig heavy chain V r
28	25	89.3	136	2	S35759	BHD9D10 protein -
29	25	89.3	138	1	D69979	conserved hypothet
30	25	89.3	145	2	D69383	conserved hypothet
31	25	89.3	154	2	B61027	hypothetical prote
32	25	89.3	171	2	S69895	helicase (EC 3.6.1
33	25	89.3	188	2	T48671	extracellular heme
34	25	89.3	203	2	C85288	hypothetical prote
35	25	89.3	203	2	T05519	hypothetical prote
36	25	89.3	220	2	H81048	biopolymer transpo
37	25	89.3	233	1	F69178	conserved hypothet
38	25	89.3	238	2	H70734	hypothetical prote
39	25	89.3	250	2	C85040	hypothetical prote
40	25	89.3	251	2	D69861	RNA polymerase sig
41	25	89.3	264	2	AF3550	phosphoglycolate p
42	25	89.3	277	2	S76396	hypothetical prote
43	25	89.3	289	2	AI3166	hypothetical prote
44	25	89.3	291	2	A69545	mRNA 3'-end proces
45	25	89.3	298	2	T29685	hypothetical prote

ALIGNMENTS

RESULT 1
S26928
Ig heavy chain V region (DP-32) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C;Accession: S26928
R;Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A;Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A;Reference number: S26885; MUID:93021117; PMID:1404388
A;Accession: S26928
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-98 <TOM>
A;Cross-references: EMBL:Z12334; NID:g32887; PIDN:CAA78204.1; PID:g32888
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 28; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 31 DYGMS 35

RESULT 2
PH1654
Ig heavy chain V region (clone 6H9) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996
C;Accession: PH1654
R;Hillson, J.L.; Karr, N.S.; Oppliger, I.R.; Mannik, M.; Sasso, E.H.
J. Exp. Med. 178, 331-336, 1993
A;Title: The structural basis of germline-encoded VH3 immunoglobulin binding to staphyl
A;Reference number: PH1642; MUID:93301610; PMID:8315388
A;Accession: PH1654
A;Molecule type: mRNA
A;Residues: 1-112 <HIL>
A;Experimental source: B cell
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;7-90/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 28; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||

```
Db         23 DYGMS 27


RESULT 3
A43413
Ig heavy chain V region - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 16-Aug-1996
C;Accession: A43413
R;Tomiyama, Y.; Brojer, E.; Ruggeri, Z.M.; Shattil, S.J.; Smiltneck, J.; Gorski, J.; Kum
J. Biol. Chem. 267, 18085-18092, 1992
A;Title: A molecular model of RGD ligands. Antibody D gene segments that direct specific
A;Reference number: A43413; MUID:92388177; PMID:1517241
A;Accession: A43413
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-119 <TOM>
A;Note: sequence extracted from NCBI backbone (NCBIP:112815)
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;9-92/Domain: immunoglobulin homology <IMM>

  Query Match             100.0%;  Score 28;  DB 2;  Length 119;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         25 DYGMS 29


RESULT 4
AB1076
probable fimbrial chaperone protein sthB [imported] - Salmonella enterica subsp. enteric
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AB1076
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AB1076
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-227 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD03425.1; PID:g16505691; GSPDB:GN00176
C;Genetics:
A;Gene: sthB
C;Superfamily: chaperone protein papD

  Query Match             100.0%;  Score 28;  DB 2;  Length 227;
  Best Local Similarity   100.0%;  Pred. No. 45;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        215 DYGMS 219


RESULT 5
H84684
En/Spm-like transposon protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: H84684
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84684
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-268 <STO>
A;Cross-references: GB:AE002093; NID:g4432833; PIDN:AAD20682.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g28440
A;Map position: 2

  Query Match             100.0%;  Score 28;  DB 2;  Length 268;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        218 DYGMS 222


RESULT 6
T04660
hypothetical protein F8D20.60 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-May-1999
C;Accession: T04660
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Jesse, T.; Heijnen, L.; Vos, P.; Mewes,
submitted to the Protein Sequence Database, July 1998
A;Reference number: Z15381
A;Accession: T04660
A;Molecule type: DNA
A;Residues: 1-268 <BEV>
A;Cross-references: EMBL:AL031135
A;Experimental source: cultivar Columbia; BAC clone F8D20
C;Genetics:
A;Map position: 4
A;Introns: 75/1; 241/1
A;Note: F8D20.60

  Query Match             100.0%;  Score 28;  DB 2;  Length 268;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        263 DYGMS 267


RESULT 7
H95307
arginine deiminase (EC 3.5.3.6) ArcB [imported] - Sinorhizobium meliloti (strain 1021) m
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
C;Accession: H95307
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: H95307
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-334 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK65026.1; PID:g14523456; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
```

A;Contents: annotation
C;Genetics:
A;Gene: arcB
A;Genome: plasmid
C;Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase
C;Keywords: hydrolase

Query Match 100.0%; Score 28; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 292 DYGMS 296

RESULT 8
H72681
hypothetical protein APE0874 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C;Accession: H72681
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: H72681
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-340 <KAW>
A;Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA79856.1; PID:g5104541
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0874
C;Superfamily: Aeropyrum pernix hypothetical protein APE0874

Query Match 100.0%; Score 28; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 82 DYGMS 86

RESULT 9
S74760
hypothetical protein slr1617 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C;Accession: S74760
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74760
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-411 <KAN>
A;Cross-references: EMBL:D90901; GB:AB001339; NID:g1651897; PIDN:BAA16911.1; PID:d101764
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 100.0%; Score 28; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 65 DYGMS 69

RESULT 10
AB3043
oxidoreductase Atu3958 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AB3043
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClel
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AB3043
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-448 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL44760.1; PID:g17742396; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu3958
A;Map position: linear chromosome

Query Match 100.0%; Score 28; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 124 DYGMS 128

RESULT 11
B98243
hypothetical protein AGR_L_1791 [imported] - Agrobacterium tumefaciens (strain C58, Cer
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C;Accession: B98243
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: B98243
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-468 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89468.1; PID:g15159336; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_1791
A;Map position: linear chromosome

Query Match 100.0%; Score 28; DB 2; Length 468;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYGMS 5
|||||
Db 144 DYGMS 148

RESULT 12
T38737
probable CAAX prenyl proteinase 1 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 07-Dec-1999
C;Accession: T38737
R;Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z21808

```
A;Accession: T38737
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-474 <GEN>
A;Cross-references: EMBL:Z68144; PIDN:CAA92258.1; GSPDB:GN00066; SPDB:SPAC3H1.05
A;Experimental source: strain 972h-; cosmid c3H1
C;Genetics:
A;Gene: SPDB:SPAC3H1.05
A;Map position: 1
A;Introns: 35/3

  Query Match             100.0%;  Score 28;  DB 2;  Length 474;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        349 DYGMS 353


RESULT 13
F89651
protein T04F8.2 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001
C;Accession: F89651
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: F89651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-478 <STO>
A;Cross-references: GB:chr_X; PIDN:CAA91478.1; PID:g3924836; GSPDB:GN00028; CESP:T04F8.2
A;Note: cDNA EST EMBL:M89094 comes from this gene
C;Genetics:
A;Gene: T04F8.2
A;Map position: X
C;Superfamily: Caenorhabditis elegans hypothetical protein T04F8.2

  Query Match             100.0%;  Score 28;  DB 2;  Length 478;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         77 DYGMS 81


RESULT 14
T24471
hypothetical protein T04F8.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T24471
R;Lennard, N.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19895
A;Accession: T24471
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-502 <WIL>
A;Cross-references: EMBL:Z66565; PIDN:CAA91478.2; GSPDB:GN00028; CESP:T04F8.2
A;Experimental source: clone T04F8
C;Genetics:
A;Gene: CESP:T04F8.2
A;Map position: X
A;Introns: 26/1; 72/1; 103/3; 162/2; 214/3; 249/2; 290/2; 340/3; 382/2; 427/1; 459/2
C;Superfamily: Caenorhabditis elegans hypothetical protein T04F8.2

  Query Match             100.0%;  Score 28;  DB 2;  Length 502;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         93 DYGMS 97


RESULT 15
H71285
probable cell division protein (ftsH) - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 21-Jan-2000
C;Accession: H71285
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: H71285
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-609 <COL>
A;Cross-references: GB:AE001247; GB:AE000520; NID:g3323059; PIDN:AAC65728.1; PID:g332306
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0765
C;Superfamily: cell division protein ftsH; FtsH/SEC18/CDC48-type ATP-binding domain homo
F;155-363/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT>

  Query Match             100.0%;  Score 28;  DB 2;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        493 DYGMS 497


Search completed: September 18, 2004, 03:47:32
Job time : 12.1786 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 02:29:47 ; Search time 7.32143 Seconds
(without alignments)
35.560 Million cell updates/sec

Title: US-10-029-926B-114
Perfect score: 28
Sequence: 1 DYGMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	100.0	334	1	OTCC_RHIME	Q92zt1 rhizobium m
2	28	100.0	354	1	CKR5_MOUSE	P51682 mus musculu
3	28	100.0	474	1	ST24_SCHPO	Q10071 schizosacch
4	28	100.0	493	1	OCLN_XENLA	Q9pun1 xenopus lae
5	28	100.0	609	1	FTSH_TREPA	O83746 treponema p
6	28	100.0	830	1	PRLR_COLLI	Q90374 columba liv
7	28	100.0	885	1	PLSB_XANAC	Q8pes0 xanthomonas
8	28	100.0	1699	1	POLN_LORDV	P54634 lordsdale v
9	28	100.0	2133	1	FA8_PIG	P12263 sus scrofa
10	28	100.0	2319	1	FA8_MOUSE	Q06194 mus musculu
11	28	100.0	2351	1	FA8_HUMAN	P00451 homo sapien
12	25	89.3	138	1	RUVX_BACSU	O34634 bacillus su
13	25	89.3	154	1	YMA6_VICFA	Q04655 vicia faba
14	25	89.3	188	1	HASA_SERMA	Q54450 serratia ma
15	25	89.3	220	1	EXBB_NEIGO	O06433 neisseria g
16	25	89.3	220	1	EXBB_NEIMA	P57027 neisseria m
17	25	89.3	220	1	EXBB_NEIMC	P95375 neisseria m
18	25	89.3	238	1	YN08_MYCTU	Q50657 mycobacteri
19	25	89.3	311	1	PPAC_STRPY	Q9a1a2 streptococc
20	25	89.3	328	1	GALE_VIBCH	Q56623 vibrio chol
21	25	89.3	333	1	NDPA_PSERE	Q8ghv6 pseudomonas
22	25	89.3	333	1	NDPA_PSESM	Q887m6 pseudomonas
23	25	89.3	334	1	NDPA_PSEAE	Q9hxf7 pseudomonas
24	25	89.3	335	1	NDPA_PSEPK	Q88p80 pseudomonas
25	25	89.3	344	1	TRPD_XYLFA	Q9pgt6 xylella fas
26	25	89.3	344	1	TRPD_XYLFT	Q87ex3 xylella fas
27	25	89.3	379	1	NER2_RAT	Q64627 rattus norv
28	25	89.3	416	1	RHL_MACFA	Q28481 macaca fasc
29	25	89.3	416	1	RHL_MACMU	Q28849 macaca mula
30	25	89.3	437	1	ARGA_HAEDU	Q7vln8 haemophilus
31	25	89.3	477	1	NIFE_RHOCA	P19055 rhodobacter
32	25	89.3	489	1	DPY1_CAEEL	Q21773 caenorhabdi
33	25	89.3	494	1	CROM_OMMSL	P30842 ommastrephe
34	25	89.3	495	1	CROM_OCTDO	P30841 octopus dof
35	25	89.3	503	1	VRK2_MOUSE	Q8bn21 mus musculu
36	25	89.3	508	1	VRK2_HUMAN	Q86y07 homo sapien
37	25	89.3	512	1	FUS1_YEAST	P11710 saccharomyc
38	25	89.3	520	1	DPY2_CAEEL	Q18677 caenorhabdi
39	25	89.3	557	1	FMO4_HUMAN	P31512 homo sapien
40	25	89.3	575	1	GGT_PSESP	P36267 pseudomonas
41	25	89.3	619	1	GATE_METTH	O26803 methanobact
42	25	89.3	761	1	SEPR_MOUSE	P97321 mus musculu
43	25	89.3	880	1	DPO1_BACSU	O34996 bacillus su
44	25	89.3	881	1	HELI_HSVEB	P28934 equine herp
45	25	89.3	881	1	HELI_VZVD	P09303 varicella-z

ALIGNMENTS

RESULT 1
OTCC_RHIME
ID OTCC_RHIME STANDARD; PRT; 334 AA.
AC Q92ZT1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ornithine carbamoyltransferase, catabolic (EC 2.1.3.3) (OTCase).
GN ARCB OR RA0368 OR SMA0695.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
CC + L-citrulline.
CC -!- PATHWAY: Arginine degradation; arginine deiminase pathway; second
CC step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the ATCase/OTCase family.
CC ----------

CC ----------
DR EMBL; AE007228; AAK65026.1; -.
DR PIR; H95307; H95307.
DR HAMAP; MF_01109; -; 1.
DR InterPro; IPR006130; Asp/Orn_COtranf.
DR InterPro; IPR002292; Orn_carbtransf.
DR InterPro; IPR006131; OTCace_O.
DR InterPro; IPR006132; OTCace_P.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace_N; 1.
DR PRINTS; PR00100; AOTCASE.
DR TIGRFAMs; TIGR00658; orni_carb_tr; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
KW Arginine metabolism; Transferase; Plasmid; Complete proteome.
FT SITE 32 32 IMPORTANT FOR STRUCTURAL INTEGRITY (BY
FT SIMILARITY).

```
FT   SITE         57     61       CARBAMOYLPHOSPHATE BINDING (BY
FT                                 SIMILARITY).
FT   SITE        108    108       CARBAMOYLPHOSPHATE BINDING (BY
FT                                 SIMILARITY).
FT   SITE        135    135       CARBAMOYLPHOSPHATE BINDING (BY
FT                                 SIMILARITY).
FT   SITE        148    148       IMPORTANT FOR STRUCTURAL INTEGRITY (BY
FT                                 SIMILARITY).
FT   SITE        273    276       ORNITHINE BINDING (BY SIMILARITY).
SQ   SEQUENCE   334 AA;  37659 MW;  BD7A703ABDB63DAE CRC64;

  Query Match             100.0%;  Score 28;  DB 1;  Length 334;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        292 DYGMS 296


RESULT 2
CKR5_MOUSE
ID   CKR5_MOUSE     STANDARD;      PRT;   354 AA.
AC   P51682; O35313; O35891; P97308; P97405; Q61867;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (MIP-1
DE   alpha receptor).
GN   CCR5 OR CMKBR5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ; TISSUE=Spleen;
RX   MEDLINE=96205938; PubMed=8631787;
RA   Boring L., Gosling J., Monteclaro F.S., Lusis A.J., Tsou C.-L.,
RA   Charo I.F.;
RT   "Molecular cloning and functional expression of murine JE (monocyte
RT   chemoattractant protein 1) and murine macrophage inflammatory protein
RT   1alpha receptors: evidence for two closely linked C-C chemokine
RT   receptors on chromosome 9.";
RL   J. Biol. Chem. 271:7551-7558(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6 X CBA; TISSUE=Thymus;
RX   MEDLINE=96278910; PubMed=8662890;
RA   Meyer A., Coyle A.J., Proudfoot A.E.I., Wells T.N.C., Power C.A.;
RT   "Cloning and characterization of a novel murine macrophage
RT   inflammatory protein-1 alpha receptor.";
RL   J. Biol. Chem. 271:14445-14451(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Ola;
RA   Kuziel W.A., Beck M.A., Dawson T.C., Maeda N.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6, and NIH Swiss; TISSUE=Kidney, Liver, and Spleen;
RX   MEDLINE=98001387; PubMed=9343222;
RA   Kuhmann S.E., Platt E.J., Kozak S.L., Kabat D.;
RT   "Polymorphisms in the CCR5 genes of African green monkeys and mice
RT   implicate specific amino acids in infections by simian and human
RT   immunodeficiency viruses.";
RL   J. Virol. 71:8642-8656(1997).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129;
RX   MEDLINE=97404635; PubMed=9261347;
RA   Doranz B.J., Lu Z.H., Rucker J., Zhang T.Y., Sharron M., Cen Y.H.,
RA   Wang Z.X., Guo H.H., Du J.G., Accavitti M.A., Doms R.W., Peiper S.C.;
RT   "Two distinct CCR5 domains can mediate coreceptor usage by human
RT   immunodeficiency virus type 1.";
RL   J. Virol. 71:6305-6314(1997).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Guo B., Kuno K., Harada A., Matsushima K.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for a C-C type chemokine. Binds to MIP-1-alpha,
CC       MIP-1-beta and RANTES and subsequently transduces a signal by
CC       increasing the intracellular calcium ions level.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Detected in monocyte/macrophage cell lines,
CC       but not in nonhematopoietic cell lines.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U47036; AAC52454.1; -.
DR   EMBL; X94151; CAA63867.1; -.
DR   EMBL; U68565; AAB37273.1; -.
DR   EMBL; U83327; AAC53386.1; -.
DR   EMBL; AF022990; AAC53389.1; -.
DR   EMBL; AF019772; AAB71183.1; -.
DR   EMBL; D83648; BAA12024.1; -.
DR   MGD; MGI:107182; Ccr5.
DR   GO; GO:0016493; F:C-C chemokine receptor activity; IDA.
DR   GO; GO:0006952; P:defense response; IMP.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Polymorphism.
FT   DOMAIN       1     32       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    33     60       1 (POTENTIAL).
FT   DOMAIN      61     70       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    71     91       2 (POTENTIAL).
FT   DOMAIN      92    104       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   105    126       3 (POTENTIAL).
FT   DOMAIN     127    143       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   144    168       4 (POTENTIAL).
FT   DOMAIN     169    200       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   201    220       5 (POTENTIAL).
FT   DOMAIN     221    237       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   238    262       6 (POTENTIAL).
FT   DOMAIN     263    279       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   280    303       7 (POTENTIAL).
FT   DOMAIN     304    354       CYTOPLASMIC (POTENTIAL).
FT   DISULFID   103    180       BY SIMILARITY.
FT   CARBOHYD   270    270       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     11     11       I -> S.
FT   VARIANT     62     62       K -> R.
FT   VARIANT     66     66       V -> M.
FT   VARIANT     97     97       I -> V.
FT   VARIANT    109    109       V -> L.
FT   VARIANT    156    156       V -> A.
FT   VARIANT    160    160       F -> S.
FT   VARIANT    185    185       P -> L.
FT   VARIANT    213    213       I -> V.
FT   VARIANT    318    318       I -> M.
FT   VARIANT    337    337       V -> A.
FT   CONFLICT     3      3       F -> L (IN REF. 2).
FT   CONFLICT    80     80       L -> F (IN REF. 2).
FT   CONFLICT   145    145       N -> I (IN REF. 5).
FT   CONFLICT   190    190       H -> Y (IN REF. 3).
FT   CONFLICT   208    208       P -> S (IN REF. 1).
SQ   SEQUENCE   354 AA;  40863 MW;  B4A6B942E88F9CF0 CRC64;
```

```
  Query Match             100.0%;  Score 28;  DB 1;  Length 354;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         15 DYGMS 19


RESULT 3
ST24_SCHPO
ID   ST24_SCHPO     STANDARD;      PRT;   474 AA.
AC   Q10071;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Probable CAAX prenyl protease 1 (EC 3.4.24.84) (Prenyl protein-
DE   specific endoprotease 1) (PPSEP 1).
GN   SPAC3H1.05.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- FUNCTION: Proteolytically removes the C-terminal three residues of
CC       farnesylated proteins (By similarity).
CC   -!- CATALYTIC ACTIVITY: The peptide bond hydrolyzed can be designated
CC       -C-|-AAX in which C is an S-isoprenylated cysteine residue, A is
CC       usually aliphatic and X is the C-terminal residue of the substrate
CC       protein, and may be any of several amino acids.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC       reticulum (By similarity).
CC   -!- SIMILARITY: Belongs to peptidase family M48.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z68144; CAA92258.1; -.
DR   PIR; T38737; T38737.
DR   MEROPS; M48.001; -.
DR   GeneDB_SPombe; SPAC3H1.05; -.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR001915; Peptidase_M48.
DR   Pfam; PF01435; Peptidase_M48; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Hypothetical protein; Hydrolase; Metalloprotease; Zinc; Transmembrane;
KW   Endoplasmic reticulum.
FT   TRANSMEM     103     123       POTENTIAL.
FT   TRANSMEM     196     216       POTENTIAL.
FT   TRANSMEM     230     250       POTENTIAL.
FT   TRANSMEM     344     364       POTENTIAL.
FT   TRANSMEM     381     401       POTENTIAL.
FT   METAL        332     332       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE     333     333       BY SIMILARITY.
FT   METAL        336     336       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL        411     411       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE     415     415       PROTON DONOR (BY SIMILARITY).
SQ   SEQUENCE   474 AA;  54001 MW;  50AC2F96E25C76C5 CRC64;

  Query Match             100.0%;  Score 28;  DB 1;  Length 474;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        349 DYGMS 353


RESULT 4
OCLN_XENLA
ID   OCLN_XENLA     STANDARD;      PRT;   493 AA.
AC   Q9PUN1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Occludin.
GN   OCLN.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A., AND PHOSPHORYLATION OF THR-375 AND SER-379.
RC   TISSUE=Ovary;
RX   MEDLINE=99421641; PubMed=10491082;
RA   Cordenonsi M., Turco F., D'Atri F., Hammar E., Martinucci G.,
RA   Meggio F., Citi S.;
RT   "Xenopus laevis occludin. Identification of in vitro phosphorylation
RT   sites by protein kinase CK2 and association with cingulin.";
RL   Eur. J. Biochem. 264:374-384(1999).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=98034414; PubMed=9365283;
RA   Cordenonsi M., Mazzon E., De Rigo L., Baraldo S., Meggio F., Citi S.;
RT   "Occludin dephosphorylation in early development of Xenopus laevis.";
RL   J. Cell Sci. 110:3131-3139(1997).
CC   -!- FUNCTION: Probably plays a role in the formation and regulation of
CC       the tight junction (TJ) paracellular permeability barrier.
CC   -!- SUBUNIT: Interacts in vitro with cingulin, possibly directly.
CC       Interacts with ZO-1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Localized at tight junctions of both
CC       epithelial and endothelial cells.
CC   -!- DEVELOPMENTAL STAGE: A maternally synthesized protein. Found in
CC       granules in the peripheral cytoplasm in the fertilized egg, it
CC       localizes first to the basolateral membrane, then to tight
CC       junctions after cingulin and ZO-1. Nascent tight junctions are in
CC       place by the two-cell stage. The maternal form is more highly
```

```
CC       phosphorylated than the form detected in later developmental
CC       stages.
CC   -!- DOMAIN: The C-terminus is cytoplasmic and is important for
CC       interaction with ZO-1. Necessary for the tight junction
CC       localization. Involved in the regulation of the permeability
CC       barrier function of the tight junction (By similarity).
CC   -!- PTM: Phosphorylated.
CC   -!- SIMILARITY: Belongs to the ELL / occludin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF170275; AAD53725.1; -.
DR   InterPro; IPR008253; Marvel.
DR   InterPro; IPR002958; Occludin.
DR   Pfam; PF01284; MARVEL; 1.
DR   PRINTS; PR01258; OCCLUDIN.
KW   Tight junction; Transmembrane; Coiled coil; Phosphorylation.
FT   DOMAIN        1     47       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     48     70       POTENTIAL.
FT   DOMAIN       71    116       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    117    141       POTENTIAL.
FT   DOMAIN      142    151       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    152    176       POTENTIAL.
FT   DOMAIN      177    224       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    225    246       POTENTIAL.
FT   DOMAIN      247    493       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      396    428       COILED COIL (POTENTIAL).
FT   MOD_RES     375    375       PHOSPHORYLATION (BY CK2) (IN VITRO).
FT   MOD_RES     379    379       PHOSPHORYLATION (BY CK2) (IN VITRO).
SQ   SEQUENCE   493 AA;  55111 MW;  9694CD302BEBBEDE CRC64;

  Query Match             100.0%;  Score 28;  DB 1;  Length 493;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        333 DYGMS 337


RESULT 5
FTSH_TREPA
ID   FTSH_TREPA     STANDARD;      PRT;   609 AA.
AC   O83746;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Cell division protein ftsH homolog (EC 3.4.24.-).
GN   FTSH OR TP0765.
OS   Treponema pallidum.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX   NCBI_TaxID=160;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nichols;
RX   MEDLINE=98332770; PubMed=9665876;
RA   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA   Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA   Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA   Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA   McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA   Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA   Venter J.C.;
RT   "Complete genome sequence of Treponema pallidum, the syphilis
RT   spirochete.";
RL   Science 281:375-388(1998).
CC   -!- FUNCTION: Seems to act as an ATP-dependent zinc metallopeptidase
CC       (By similarity).
CC   -!- COFACTOR: Binds 1 zinc ion (Potential).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the AAA ATPase family.
CC   -!- SIMILARITY: Belongs to peptidase family M41.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001247; AAC65728.1; -.
DR   PIR; H71285; H71285.
DR   TIGR; TP0765; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003959; AAA_ATPase_centr.
DR   InterPro; IPR003960; AAA_sub.
DR   InterPro; IPR005936; Peptidase_FtsH.
DR   InterPro; IPR000642; Peptidase_M41.
DR   Pfam; PF00004; AAA; 1.
DR   Pfam; PF01434; Peptidase_M41; 1.
DR   SMART; SM00382; AAA; 1.
DR   TIGRFAMs; TIGR01241; FtsH_fam; 1.
DR   PROSITE; PS00674; AAA; 1.
KW   Cell division; ATP-binding; Transmembrane; Hydrolase; Metalloprotease;
KW   Zinc; Complete proteome.
FT   TRANSMEM     89    106       POTENTIAL.
FT   NP_BIND     181    188       ATP (POTENTIAL).
FT   METAL       402    402       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    403    403       BY SIMILARITY.
FT   METAL       406    406       ZINC (CATALYTIC) (BY SIMILARITY).
SQ   SEQUENCE   609 AA;  67544 MW;  0AE932C2A4F6A5D4 CRC64;

  Query Match             100.0%;  Score 28;  DB 1;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 61;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        493 DYGMS 497


RESULT 6
PRLR_COLLI
ID   PRLR_COLLI     STANDARD;      PRT;   830 AA.
AC   Q90374;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Prolactin receptor precursor (PRL-R).
GN   PRLR.
OS   Columba livia (Domestic pigeon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
OX   NCBI_TaxID=8932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cropsac;
RX   MEDLINE=94283267; PubMed=7516866;
RA   Chen X., Horseman N.D.;
RT   "Cloning, expression, and mutational analysis of the pigeon prolactin
RT   receptor.";
RL   Endocrinology 135:269-276(1994).
CC   -!- FUNCTION: This is a receptor for the anterior pituitary hormone
CC       prolactin.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the type I cytokine family of receptors.
CC       Subfamily 1.
CC   -!- SIMILARITY: Contains 4 fibronectin type III domains.
CC   ---------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U07694; AAA20646.1; -.
DR   PIR; I50455; I50455.
DR   HSSP; P16471; 1BP3.
DR   InterPro; IPR002996; CR1A.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR003528; Hemtopoptn_L_F1.
DR   Pfam; PF00041; fn3; 4.
DR   SMART; SM00060; FN3; 4.
DR   PROSITE; PS01352; HEMATOPO_REC_L_F1; 1.
KW   Receptor; Transmembrane; Glycoprotein; Signal; Repeat.
FT   SIGNAL       1     23       POTENTIAL.
FT   CHAIN        24    830      PROLACTIN RECEPTOR.
FT   DOMAIN       24    439      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     440   460      POTENTIAL.
FT   DOMAIN       461   830      CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       25    122      FIBRONECTIN TYPE-III 1.
FT   DOMAIN       123   226      FIBRONECTIN TYPE-III 2.
FT   DOMAIN       229   326      FIBRONECTIN TYPE-III 3.
FT   DOMAIN       327   429      FIBRONECTIN TYPE-III 4.
FT   DISULFID     36    46       BY SIMILARITY.
FT   DISULFID     75    86       BY SIMILARITY.
FT   CARBOHYD     59    59       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     91    91       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     100   100      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     112   112      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     132   132      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     263   263      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     304   304      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     316   316      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     336   336      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   830 AA;  94507 MW;  3B074E83CDF69EFF CRC64;

  Query Match             100.0%;  Score 28;  DB 1;  Length 830;
  Best Local Similarity   100.0%;  Pred. No. 82;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         77 DYGMS 81


RESULT 7
PLSB_XANAC
ID   PLSB_XANAC     STANDARD;      PRT;   885 AA.
AC   Q8PES0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).
GN   PLSB OR XAC4270.
OS   Xanthomonas axonopodis (pv. citri).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=92829;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=306 / ATCC 13902 / XV 101;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
CC   -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC       acyl-sn-glycerol 3-phosphate.
CC   -!- PATHWAY: De novo phospholipid biosynthesis; first step. May also
CC       function in the regulation of membrane biogenesis.
CC   -!- SUBCELLULAR LOCATION: Membrane-bound (By similarity).
CC   -!- SIMILARITY: Belongs to the GPAT / DAPAT family.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE012079; AAM39105.1; -.
DR   HAMAP; MF_00393; -; 1.
DR   InterPro; IPR002123; Acyltransferase.
DR   Pfam; PF01553; Acyltransferase; 1.
KW   Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;
KW   Complete proteome.
SQ   SEQUENCE   885 AA;  98467 MW;  FC5942DBAB3B9825 CRC64;

  Query Match             100.0%;  Score 28;  DB 1;  Length 885;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         96 DYGMS 100


RESULT 8
POLN_LORDV
ID   POLN_LORDV     STANDARD;      PRT;  1699 AA.
AC   P54634;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Non-structural polyprotein [Contains: RNA-directed RNA polymerase
DE   (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like
DE   protein)].
OS   Lordsdale virus (Human enteric calicivirus).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC   Norovirus.
OX   NCBI_TaxID=82658;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96005060; PubMed=7561776;
RA   Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;
RT   "Human enteric Caliciviridae: the complete genome sequence and
RT   expression of virus-like particles from a genetic group II small
RT   round structured virus.";
RL   J. Gen. Virol. 76:2349-2355(1995).
CC   -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).
CC   -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR
CC       THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY
CC       SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
```

```
CC   -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.
CC   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X86557; CAA60254.1; -.
DR   MEROPS; C37.001; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR004004; Calici_pol_hel.
DR   InterPro; IPR001665; Peptidase_C37.
DR   InterPro; IPR000605; RNA_helicase.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF05416; Peptidase_C37; 1.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   Pfam; PF00910; RNA_helicase; 1.
DR   PRINTS; PR00918; CALICVIRUSNS.
DR   PRINTS; PR00917; SRSVCYSPTASE.
DR   SMART; SM00382; AAA; 1.
KW   Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW   Thiol protease; Helicase; ATP-binding.
FT   CHAIN          ?      ?       HELICASE (P2C-LIKE).
FT   CHAIN      ?1143  ?1182       PROTEASE.
FT   NP_BIND     495    502       ATP (POTENTIAL).
FT   ACT_SITE   1147   1147       PROTEASE (BY SIMILARITY).
FT   ACT_SITE   1165   1165       PROTEASE (BY SIMILARITY).
SQ   SEQUENCE   1699 AA;  189199 MW;  FA00B3B67FF3A0B6 CRC64;

  Query Match             100.0%;  Score 28;  DB 1;  Length 1699;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        546 DYGMS 550


RESULT 9
FA8_PIG
ID   FA8_PIG        STANDARD;      PRT;  2133 AA.
AC   P12263; Q95243;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor VIII precursor (Procoagulant component).
GN   F8 OR CF8.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Healey J.F., Lubin I.M., Lollar P.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 705-1573 FROM N.A.
RX   MEDLINE=86287369; PubMed=3016730;
RA   Toole J.J., Pittman D.D., Orr E.C., Murtha P., Wasley L.C.,
RA   Kaufman R.J.;
RT   "A large region (approximately equal to 95 kDa) of human factor VIII
RT   is dispensable for in vitro procoagulant activity.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5939-5942(1986).
RN   [3]
RP   SEQUENCE OF 392-759 FROM N.A.
RX   MEDLINE=94179260; PubMed=7510693;
RA   Lubin I.M., Healey J.F., Scandella D., Runge M.S., Lollar P.;
RT   "Elimination of a major inhibitor epitope in factor VIII.";
RL   J. Biol. Chem. 269:8639-8641(1994).
CC   -!- FUNCTION: Factor VIII, along with calcium and phospholipid, acts
CC       as a cofactor for factor IXa when it converts factor X to the
CC       activated form, factor Xa.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- SIMILARITY: Contains 3 F5/8 type A domains.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC   -!- SIMILARITY: STRONG, TO COAGULATION FACTOR V.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U49517; AAB06705.1; -.
DR   PIR; A25945; A25945.
DR   PIR; T42763; T42763.
DR   HSSP; P00451; 1CFG.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00394; Cu-oxidase; 3.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
KW   Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
KW   Signal; Glycoprotein; Sulfation.
FT   SIGNAL         1     19       POTENTIAL.
FT   CHAIN         20   2133       COAGULATION FACTOR VIII.
FT   DOMAIN        20    357       F5/8 TYPE A 1.
FT   DOMAIN        20    199       PLASTOCYANIN-LIKE 1.
FT   DOMAIN       207    357       PLASTOCYANIN-LIKE 2.
FT   DOMAIN       399    730       F5/8 TYPE A 2.
FT   DOMAIN       399    573       PLASTOCYANIN-LIKE 3.
FT   DOMAIN       583    730       PLASTOCYANIN-LIKE 4.
FT   DOMAIN       760   1599       B.
FT   DOMAIN      1495   1822       F5/8 TYPE A 3.
FT   DOMAIN      1495   1659       PLASTOCYANIN-LIKE 5.
FT   DOMAIN      1669   1822       PLASTOCYANIN-LIKE 6.
FT   DOMAIN      1822   1970       F5/8 TYPE C 1.
FT   DOMAIN      1975   2127       F5/8 TYPE C 2.
FT   SITE         391    392       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   SITE         759    760       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   SITE        1449   1450       CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT   SITE        1490   1491       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   MOD_RES      737    737       SULFATION (BY SIMILARITY).
FT   MOD_RES      738    738       SULFATION (BY SIMILARITY).
FT   MOD_RES      742    742       SULFATION (BY SIMILARITY).
FT   DISULFID     173    199       PROBABLE.
FT   DISULFID     547    573       PROBABLE.
FT   DISULFID    1633   1659       PROBABLE.
FT   DISULFID    1822   1970       BY SIMILARITY.
FT   DISULFID    1975   2127       BY SIMILARITY.
FT   CARBOHYD     233    233       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     259    259       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     601    601       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     929    929       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     985    985       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1025   1025       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1111   1111       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1181   1181       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1208   1208       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1245   1245       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1265   1265       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1335   1335       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD   1408   1408       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1611   1611       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1919   1919       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    713    713       N -> M (IN REF. 2).
FT   CONFLICT    734    734       I -> T (IN REF. 2).
FT   CONFLICT    792    792       G -> Q (IN REF. 2).
FT   CONFLICT   1133   1133       E -> F (IN REF. 2).
FT   CONFLICT   1191   1191       I -> L (IN REF. 2).
FT   CONFLICT   1209   1209       R -> F (IN REF. 2).
FT   CONFLICT   1437   1437       C -> G (IN REF. 2).
FT   CONFLICT   1456   1456       F -> R (IN REF. 2).
FT   CONFLICT   1539   1539       F -> R (IN REF. 2).
FT   CONFLICT   1546   1546       Q -> N (IN REF. 2).
SQ   SEQUENCE   2133 AA;  239304 MW;  152BBA8997F570DA CRC64;

  Query Match             100.0%;  Score 28;  DB 1;  Length 2133;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db       1509 DYGMS 1513


RESULT 10
FA8_MOUSE
ID   FA8_MOUSE      STANDARD;      PRT;  2319 AA.
AC   Q06194;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor VIII precursor (Procoagulant component).
GN   F8 OR CF8 OR F8C.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX   MEDLINE=93300511; PubMed=8314577;
RA   Elder B., Lakich D., Gitschier J.;
RT   "Sequence of the murine factor VIII cDNA.";
RL   Genomics 16:374-379(1993).
CC   -!- FUNCTION: Factor VIII, along with calcium and phospholipid, acts
CC       as a cofactor for factor IXa when it converts factor X to the
CC       activated form, factor Xa.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: Found in most tissues.
CC   -!- SIMILARITY: Contains 3 F5/8 type A domains.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC   -!- SIMILARITY: STRONG, TO COAGULATION FACTOR V.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L05573; AAA37385.1; -.
DR   PIR; A47004; A47004.
DR   HSSP; P00451; 1CFG.
DR   MGD; MGI:88383; F8.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00394; Cu-oxidase; 3.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
KW   Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
KW   Signal; Glycoprotein; Sulfation.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20   2319       COAGULATION FACTOR VIII.
FT   DOMAIN       20    349       F5/8 TYPE A 1.
FT   DOMAIN       20    199       PLASTOCYANIN-LIKE 1.
FT   DOMAIN      207    349       PLASTOCYANIN-LIKE 2.
FT   DOMAIN      399    730       F5/8 TYPE A 2.
FT   DOMAIN      399    573       PLASTOCYANIN-LIKE 3.
FT   DOMAIN      583    730       PLASTOCYANIN-LIKE 4.
FT   DOMAIN      760   1640       B.
FT   DOMAIN     1683   2008       F5/8 TYPE A 3.
FT   DOMAIN     1683   1845       PLASTOCYANIN-LIKE 5.
FT   DOMAIN     1855   2008       PLASTOCYANIN-LIKE 6.
FT   DOMAIN     2008   2156       F5/8 TYPE C 1.
FT   DOMAIN     2161   2313       F5/8 TYPE C 2.
FT   SITE        391    392       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   SITE        759    760       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   SITE       1678   1679       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   SITE       1324   1325       CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT   SITE       1640   1641       CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT   MOD_RES     367    367       SULFATION (BY SIMILARITY).
FT   MOD_RES     737    737       SULFATION (BY SIMILARITY).
FT   MOD_RES     738    738       SULFATION (BY SIMILARITY).
FT   MOD_RES     742    742       SULFATION (BY SIMILARITY).
FT   MOD_RES    1669   1669       SULFATION (REQUIRED FOR VWF BINDING)
FT                                (BY SIMILARITY).
FT   MOD_RES    1687   1687       SULFATION (BY SIMILARITY).
FT   DISULFID    173    199       PROBABLE.
FT   DISULFID    547    573       PROBABLE.
FT   DISULFID   1819   1845       PROBABLE.
FT   DISULFID   2008   2156       BY SIMILARITY.
FT   DISULFID   2161   2313       BY SIMILARITY.
FT   CARBOHYD     61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    233    233       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    259    259       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    423    423       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    601    601       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    880    880       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    958    958       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1015   1015       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1022   1022       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1026   1026       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1044   1044       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1076   1076       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1087   1087       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1136   1136       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1161   1161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1192   1192       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1255   1255       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1268   1268       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1273   1273       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1274   1274       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1302   1302       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1316   1316       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1340   1340       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1378   1378       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1797   1797       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2105   2105       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   2319 AA;  266148 MW;  FD054DE051DB2A01 CRC64;

  Query Match             100.0%;  Score 28;  DB 1;  Length 2319;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db       1697 DYGMS 1701
```

```
RESULT 11
FA8_HUMAN
ID   FA8_HUMAN      STANDARD;      PRT;  2351 AA.
AC   P00451;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor VIII precursor (Procoagulant component)
DE   (Antihemophilic factor) (AHF).
GN   F8 OR F8C.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86081164; PubMed=3935400;
RA   Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,
RA   Hartog K., Kuo C.H., Masiarz F.R., Merryweather J.P., Najarian R.,
RA   Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,
RA   Urdea M.S., Valenzuela P., Dahl H.-H.M., Favalaro J., Hansen J.,
RA   Nordfang O., Ezban M.;
RT   "Characterization of the polypeptide composition of human factor
RT   VIII:C and the nucleotide sequence and expression of the human kidney
RT   cDNA.";
RL   DNA 4:333-349(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85061548; PubMed=6438526;
RA   Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,
RA   Keyt B., Seeburg P.H., Smith D.H., Hollingshead P., Wion K.L.,
RA   Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;
RT   "Expression of active human factor VIII from recombinant DNA clones.";
RL   Nature 312:330-337(1984).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85061550; PubMed=6438528;
RA   Toole J.J., Knopf J.L., Wozney J.M., Sultzman L.A., Buecker J.L.,
RA   Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,
RA   Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Fass D.N.,
RA   Hewick R.M.;
RT   "Molecular cloning of a cDNA encoding human antihaemophilic factor.";
RL   Nature 312:342-347(1984).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93265012; PubMed=1303178;
RA   Gitschier J., Wood W.I.;
RT   "Sequence of the exon-containing regions of the human factor VIII
RT   gene.";
RL   Hum. Mol. Genet. 1:199-200(1992).
RN   [5]
RP   SEQUENCE OF 2064-2070 FROM N.A.
RA   de Water N.S., Williams R., Browett P.J.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SULFATION OF TYR-1699.
RX   MEDLINE=91093266; PubMed=1898735;
RA   Leyte A., van Schijndel H.B., Niehrs C., Huttner W.B., Verbeet M.P.,
RA   Mertens K., van Mourik J.A.;
RT   "Sulfation of Tyr1680 of human blood coagulation factor VIII is
RT   essential for the interaction of factor VIII with von Willebrand
RT   factor.";
RL   J. Biol. Chem. 266:740-746(1991).
RN   [7]
RP   SULFATION.
RX   MEDLINE=92207952; PubMed=1554716;
RA   Pittman D.D., Wang J.H., Kaufman R.J.;
RT   "Identification and functional importance of tyrosine sulfate
RT   residues within recombinant factor VIII.";
RL   Biochemistry 31:3315-3325(1992).
RN   [8]
RP   STRUCTURE BY NMR OF 2322-2343.
RX   MEDLINE=95200924; PubMed=7893714;
RA   Gilbert G.E., Baleja J.D.;
RT   "Membrane-binding peptide from the C2 domain of factor VIII forms an
RT   amphipathic structure as determined by NMR spectroscopy.";
RL   Biochemistry 34:3022-3031(1995).
RN   [9]
RP   REVIEW ON MOLECULAR BASIS OF HEMA.
RX   MEDLINE=91221499; PubMed=1902642;
RA   Gitschier J.;
RT   "The molecular basis of hemophilia A.";
RL   Ann. N.Y. Acad. Sci. 614:89-96(1991).
RN   [10]
RP   REVIEW ON MOLECULAR BASIS OF HEMA.
RX   MEDLINE=89088506; PubMed=2491949;
RA   White G.C. II, Shoemaker C.B.;
RT   "Factor VIII gene and hemophilia A.";
RL   Blood 73:1-12(1989).
RN   [11]
RP   REVIEW ON MOLECULAR BASIS OF HEMA.
RX   MEDLINE=95245332; PubMed=7728145;
RA   Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
RT   "Molecular etiology of factor VIII deficiency in hemophilia A.";
RL   Hum. Mutat. 5:1-22(1995).
RN   [12]
RP   VARIANT HEMA GLN-2326.
RX   MEDLINE=86235434; PubMed=3012775;
RA   Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
RT   "Identification of a missense mutation in the factor VIII gene of a
RT   mild hemophiliac.";
RL   Science 232:1415-1416(1986).
RN   [13]
RP   VARIANT HEMA PRO-2135.
RX   MEDLINE=88096539; PubMed=3122181;
RA   Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;
RT   "A novel missense mutation in the factor VIII gene identified by
RT   analysis of amplified hemophilia DNA sequences.";
RL   Nucleic Acids Res. 15:9797-9805(1987).
RN   [14]
RP   VARIANT HEMA GLN-2228.
RX   MEDLINE=88191889; PubMed=2833855;
RA   Youssoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
RA   Kazazian H.H.;
RT   "Nonsense and missense mutations in hemophilia A: estimate of the
RT   relative mutation rate at CG dinuclectides.";
RL   Am. J. Hum. Genet. 42:718-725(1988).
RN   [15]
RP   VARIANT HEMA GLY-291.
RX   MEDLINE=88220354; PubMed=2835904;
RA   Youssoufian H., Wong C., Aronis S., Platokoukis H., Kazazian H.H. Jr.,
RA   Antonarakis S.E.;
RT   "Moderately severe hemophilia A resulting from Glu-->Gly substitution
RT   in exon 7 of the factor VIII gene.";
RL   Am. J. Hum. Genet. 42:867-871(1988).
RN   [16]
RP   VARIANT HEMA CYS-1708.
RX   MEDLINE=89274393; PubMed=2499363;
RA   O'Brien D.P., Tuddenham E.G.;
RT   "Purification and characterization of factor VIII 1,689-Cys: a
RT   nonfunctional cofactor occurring in a patient with severe hemophilia
RT   A.";
RL   Blood 73:2117-2122(1989).
RN   [17]
RP   VARIANT HEMA CYS-391.
RX   MEDLINE=90001543; PubMed=2506948;
RA   Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;
RT   "An arginine to cysteine amino acid substitution at a critical
RT   thrombin cleavage site in a dysfunctional factor VIII molecule.";
RL   Blood 74:1612-1617(1989).
RN   [18]
RP   VARIANT HEMA LEU-189.
RX   MEDLINE=90057680; PubMed=2510835;
RA   Chan V., Chan T.K., Tong T.M., Todd D.;
RT   "A novel missense mutation in exon 4 of the factor VIII:C gene
RT   resulting in moderately severe hemophilia A.";
```

```
RL   Blood 74:2688-2691(1989).
RN   [19]
RP   VARIANT HEMA LEU-2326.
RX   MEDLINE=89197216; PubMed=2495245;
RA   Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
RT   "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
RT   of the factor VIII gene.";
RL   Hum. Genet. 81:335-338(1989).
RN   [20]
RP   VARIANT HEMA HIS-391.
RX   MEDLINE=89264602; PubMed=2498882;
RA   Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
RA   Fujimaki M., Hoyer L.W.;
RT   "Direct characterization of factor VIII in plasma: detection of a
RT   mutation altering a thrombin cleavage site
RT   (arginine-372-->histidine).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
RN   [21]
RP   VARIANT HEMA CYS-1708.
RX   MEDLINE=90105723; PubMed=2104766;
RA   Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
RA   Phillips J.A. III, Janco R.L., Hoyer L.W.;
RT   "Characterization of a thrombin cleavage site mutation (Arg 1689 to
RT   Cys) in the factor VIII gene of two unrelated patients with
RT   cross-reacting material-positive hemophilia A.";
RL   Blood 75:384-389(1990).
RN   [22]
RP   VARIANTS HEMA GLN-2228 AND LEU-2326.
RX   MEDLINE=90123183; PubMed=2105106;
RA   Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
RA   Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
RT   "Recurrent mutations and three novel rearrangements in the factor
RT   VIII gene of hemophilia A patients of Italian descent.";
RL   Blood 75:662-670(1990).
RN   [23]
RP   VARIANT HEMA CYS-391.
RX   MEDLINE=90329422; PubMed=1973901;
RA   Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
RT   "CRM+ haemophilia A due to a missense mutation (372-->Cys) at the
RT   internal heavy chain thrombin cleavage site.";
RL   Br. J. Haematol. 75:73-77(1990).
RN   [24]
RP   VARIANTS HEMA PHE-1699 AND CYS-1708.
RX   MEDLINE=90152691; PubMed=2105906;
RA   Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
RA   Kazazian H.H., Antonarakis S.E.;
RT   "Characterization of mutations in the factor VIII gene by direct
RT   sequencing of amplified genomic DNA.";
RL   Genomics 6:65-71(1990).
RN   [25]
RP   VARIANTS HEMA CYS-1728 AND ASP-1941.
RX   MEDLINE=90169988; PubMed=2106480;
RA   Traystman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,
RA   Kazazian H.H.;
RT   "Use of denaturing gradient gel electrophoresis to detect point
RT   mutations in the factor VIII gene.";

  Query Match             100.0%;  Score 28;  DB 1;  Length 2351;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db       1727 DYGMS 1731


RESULT 12
RUVX_BACSU
ID   RUVX_BACSU     STANDARD;      PRT;   138 AA.
AC   O34634;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Putative Holliday junction resolvase (EC 3.1.-.-).
GN   YRRK OR BSU27390.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
CC   -!- FUNCTION: Could be a nuclease that resolves Holliday junction
CC       intermediates in genetic recombination.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Belongs to the yqgF HJR family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z99117; CAB14681.1; -.
DR   PIR; D69979; D69979.
DR   SubtiList; BG13792; yrrK.
DR   HAMAP; MF_00651; -; 1.
DR   InterPro; IPR005227; Cons_hypoth250.
DR   InterPro; IPR006641; YqgFC.
DR   Pfam; PF03652; UPF0081; 1.
DR   SMART; SM00732; YqgFc; 1.
DR   TIGRFAMs; TIGR00250; TIGR00250; 1.
KW   Hydrolase; Nuclease; DNA repair; DNA recombination; Complete proteome.
SQ   SEQUENCE   138 AA;  15210 MW;  39F3EA8A6B3B1C06 CRC64;

  Query Match             89.3%;  Score 25;  DB 1;  Length 138;
  Best Local Similarity   80.0%;  Pred. No. 70;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||:|
Db         39 DYGLS 43
```

```
RESULT 13
YMA6_VICFA
ID   YMA6_VICFA     STANDARD;      PRT;   154 AA.
AC   Q04655;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 17.9 kDa protein in ATP6 5'region (ORF154).
OS   Vicia faba (Broad bean).
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Vicia.
OX   NCBI_TaxID=3906;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91059426; PubMed=2245478;
RA   Macfarlane J.L., Wahleithner J.A., Wolstenholme D.R.;
RT   "A broad bean mitochondrial atp6 gene with an unusually simple, non-
RT   conserved 5' region.";
RL   Curr. Genet. 18:87-91(1990).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X54285; CAA38182.1; -.
DR   PIR; B61027; B61027.
KW   Hypothetical protein; Mitochondrion.
SQ   SEQUENCE   154 AA;  17934 MW;  304467D88CC6CE65 CRC64;

  Query Match             89.3%;  Score 25;  DB 1;  Length 154;
  Best Local Similarity   80.0%;  Pred. No. 78;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||:|
Db         44 DYGLS 48


RESULT 14
HASA_SERMA
ID   HASA_SERMA     STANDARD;      PRT;   188 AA.
AC   Q54450;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hemophore HasA (Heme acquisition system protein A).
GN   HASA.
OS   Serratia marcescens.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Serratia.
OX   NCBI_TaxID=615;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 1-11, FUNCTION, SUBCELLULAR LOCATION,
RP   AND INDUCTION.
RC   STRAIN=SM365;
RX   MEDLINE=95024064; PubMed=7937909;
RA   Letoffe S., Ghigo J.M., Wandersman C.;
RT   "Iron acquisition from heme and hemoglobin by a Serratia marcescens
RT   extracellular protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:9876-9880(1994).
RN   [2]
RP   SEQUENCE, AND SUBUNIT.
RX   MEDLINE=97332469; PubMed=9188703;
RA   Izadi N., Henry Y., Haladjian J., Goldberg M.E., Wandersman C.,
RA   Delepierre M., Lecroisey A.;
RT   "Purification and characterization of an extracellular heme-binding
RT   protein, HasA, involved in heme iron acquisition.";
RL   Biochemistry 36:7050-7057(1997).
RN   [3]
RP   FUNCTION.
RC   STRAIN=SM365;
RX   MEDLINE=97315228; PubMed=9171402;
RA   Ghigo J.M., Letoffe S., Wandersman C.;
RT   "A new type of hemophore-dependent heme acquisition system of Serratia
RT   marcescens reconstituted in Escherichia coli.";
RL   J. Bacteriol. 179:3572-3579(1997).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX   MEDLINE=99287096; PubMed=10360351;
RA   Arnoux P., Haser R., Izadi N., Lecroisey A., Delepierre M.,
RA   Wandersman C., Czjzek M.;
RT   "The crystal structure of HasA, a hemophore secreted by Serratia
RT   marcescens.";
RL   Nat. Struct. Biol. 6:516-520(1999).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (1.77 ANGSTROMS).
RX   MEDLINE=20422928; PubMed=10966573;
RA   Arnoux P., Haser R., Izadi-Pruneyre N., Lecroisey A., Czjzek M.;
RT   "Functional aspects of the heme bound hemophore HasA by structural
RT   analysis of various crystal forms.";
RL   Proteins 41:202-210(2000).
CC   -!- FUNCTION: Can bind free heme and also acquire it from hemoglobin.
CC       Conveys heme from hemoglobin to the HasR receptor which releases
CC       it into the bacterium. HasR alone can take up heme but the synergy
CC       between HasA and HasR increases heme uptake 100-fold.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- INDUCTION: By iron depletion.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X81195; CAA57068.1; -.
DR   PIR; T48671; T48671.
DR   PDB; 1B2V; 24-JUN-99.
DR   PDB; 1DK0; 27-DEC-00.
DR   PDB; 1DKH; 27-DEC-00.
KW   Heme; 3D-structure.
FT   METAL        32     32       IRON (HEME AXIAL LIGAND).
FT   METAL        75     75       IRON (HEME AXIAL LIGAND).
FT   CONFLICT    170    170       A -> Q (IN REF. 2).
SQ   SEQUENCE   188 AA;  19283 MW;  8E4ACC8DB254FC6D CRC64;

  Query Match             89.3%;  Score 25;  DB 1;  Length 188;
  Best Local Similarity   80.0%;  Pred. No. 95;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||:|
Db        157 DYGLS 161


RESULT 15
EXBB_NEIGO
ID   EXBB_NEIGO     STANDARD;      PRT;   220 AA.
AC   O06433;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Biopolymer transport exbB protein.
GN   EXBB.
OS   Neisseria gonorrhoeae.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
```

```
OX   NCBI_TaxID=485;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FA19;
RX   MEDLINE=97285757; PubMed=9140974;
RA   Biswas G.D., Anderson J.E., Sparling P.F.;
RT   "Cloning and functional characterization of Neisseria gonorrhoeae
RT   tonB, exbB and exbD genes.";
RL   Mol. Microbiol. 24:169-179(1997).
CC   -!- FUNCTION: Involved in the tonB-dependent energy-dependent
CC       transport of various receptor-bound substrates. Protects exbD from
CC       proteolytic degradation and functionally stabilizes tonB (By
CC       similarity).
CC   -!- SUBUNIT: The accessory proteins exbB and exbD seem to form a
CC       complex with tonB (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC   -!- SIMILARITY: Belongs to the exbB / tolQ family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U79563; AAC45287.1; -.
DR   InterPro; IPR002898; MotA_ExbB.
DR   Pfam; PF01618; MotA_ExbB; 1.
KW   Transport; Protein transport; Transmembrane; Inner membrane.
FT   TRANSMEM     13     33       POTENTIAL.
FT   TRANSMEM    126    146       POTENTIAL.
FT   TRANSMEM    168    188       POTENTIAL.
SQ   SEQUENCE   220 AA;  24132 MW;  0247734B36123FB5 CRC64;

  Query Match             89.3%;  Score 25;  DB 1;  Length 220;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              ||||:
Db        121 DYGMT 125


Search completed: September 18, 2004, 03:43:46
Job time : 9.32143 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 03:24:30 ; Search time 26.0714 Seconds
(without alignments)
60.510 Million cell updates/sec

Title: US-10-029-926B-114
Perfect score: 28
Sequence: 1 DYGMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	100.0	70	10	Q84Z47	Q84z47 oryza sativ
2	28	100.0	110	11	Q80YY5	Q80yy5 mus musculu
3	28	100.0	112	4	Q9HCC1	Q9hcc1 homo sapien
4	28	100.0	204	11	O08826	O08826 mus musculu
5	28	100.0	204	11	O08823	O08823 mus musculu
6	28	100.0	227	16	Q8ZJU0	Q8zju0 salmonella
7	28	100.0	227	16	Q8Z0S8	Q8z0s8 salmonella
8	28	100.0	249	16	Q8E110	Q8e110 streptococc
9	28	100.0	268	10	Q9SKM6	Q9skm6 arabidopsis
10	28	100.0	268	10	O81788	O81788 arabidopsis
11	28	100.0	293	16	Q8PFN0	Q8pfn0 xanthomonas
12	28	100.0	300	10	Q9C840	Q9c840 arabidopsis
13	28	100.0	302	10	Q9M7X8	Q9m7x8 arabidopsis
14	28	100.0	333	11	Q80YY6	Q80yy6 mus musculu
15	28	100.0	340	17	Q9YDP1	Q9ydp1 aeropyrum p
16	28	100.0	362	10	Q9LTC3	Q9ltc3 arabidopsis
17	28	100.0	406	2	Q9EYS9	Q9eys9 rhizobium s
18	28	100.0	411	16	P72895	P72895 synechocyst
19	28	100.0	439	16	Q87T65	Q87t65 vibrio para
20	28	100.0	448	16	Q880T6	Q880t6 pseudomonas
21	28	100.0	468	16	Q8U8Y0	Q8u8y0 agrobacteri
22	28	100.0	486	10	Q8L4E4	Q8l4e4 oryza sativ
23	28	100.0	502	5	Q22162	Q22162 caenorhabdi
24	28	100.0	532	10	Q7XR43	Q7xr43 oryza sativ
25	28	100.0	537	12	Q91F86	Q91f86 chilo iride
26	28	100.0	572	3	P87247	P87247 botrytis ci
27	28	100.0	572	16	Q98AA9	Q98aa9 rhizobium l
28	28	100.0	603	16	Q899H3	Q899h3 clostridium
29	28	100.0	650	16	Q9A554	Q9a554 caulobacter
30	28	100.0	728	16	Q7URM7	Q7urm7 rhodopirell
31	28	100.0	756	10	Q9M999	Q9m999 arabidopsis
32	28	100.0	785	4	Q86W01	Q86w01 homo sapien
33	28	100.0	916	10	Q9ZUF5	Q9zuf5 arabidopsis
34	28	100.0	1008	10	Q9XEC0	Q9xec0 arabidopsis
35	28	100.0	1076	4	Q8N3A9	Q8n3a9 homo sapien
36	28	100.0	1222	4	Q86T93	Q86t93 homo sapien
37	28	100.0	1236	4	Q96JH7	Q96jh7 homo sapien
38	28	100.0	1250	5	Q23409	Q23409 caenorhabdi
39	28	100.0	1699	12	Q8V0P3	Q8v0p3 human calic
40	28	100.0	1699	12	Q8JXI8	Q8jxi8 norwalk-lik
41	28	100.0	1699	12	Q68103	Q68103 hawaii cali
42	28	100.0	1699	12	Q9WI83	Q9wi83 camberwell
43	28	100.0	1699	12	Q80RD7	Q80rd7 snow mounta
44	28	100.0	1702	12	Q8JXI5	Q8jxi5 norwalk-lik
45	28	100.0	1702	12	Q8JXI4	Q8jxi4 norwalk-lik

ALIGNMENTS

RESULT 1
Q84Z47
ID Q84Z47 PRELIMINARY; PRT; 70 AA.
AC Q84Z47;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE P0458G06.5 protein.
GN P0458G06.5.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
RT clone:P0458G06.";
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP005051; BAC57333.1; -.
SQ SEQUENCE 70 AA; 7281 MW; FE752CC92EB2B855 CRC64;

Query Match 100.0%; Score 28; DB 10; Length 70;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 DYGMS 5
              |||||
Db         34 DYGMS 38
```

RESULT 2
Q80YY5
ID Q80YY5 PRELIMINARY; PRT; 110 AA.
AC Q80YY5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

```
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   BM289L21.4.2 (Novel protein, variant 2) (Fragment).
GN   BM289L21.4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   North P., Leaves N., Greystrong J., Coppola M., Manjunath S.,
RA   Russell E., Smith M., Strachan G., Tofts C., Boal E., Cobley V.,
RA   Hunter G., Kimberley C., Thomas D., Cave-Berry L., Weston P.,
RA   Botcherby M.R.M.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX247953; CAD83028.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   110 AA;  12302 MW;  B5D4F9A90768B46F CRC64;

  Query Match             100.0%;  Score 28;  DB 11;  Length 110;
  Best Local Similarity   100.0%;  Pred. No. 65;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         49 DYGMS 53


RESULT 3
Q9HCC1
ID   Q9HCC1      PRELIMINARY;      PRT;   112 AA.
AC   Q9HCC1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Single chain Fv (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kikuchi M., Takeda C., Tsujimoto Y., Asada S., Nagata K.;
RT   "An antibody fragment2A3 specific for native lysozyme :Isolaion from a
RT   human synthetic phage display library and characterization.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB049915; BAB16829.1; -.
DR   HSSP; P01772; 2FB4.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
FT   NON_TER       1      1
FT   NON_TER     112    112
SQ   SEQUENCE   112 AA;  12243 MW;  24F1A45EC3B84788 CRC64;

  Query Match             100.0%;  Score 28;  DB 4;  Length 112;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         31 DYGMS 35


RESULT 4
O08826
ID   O08826      PRELIMINARY;      PRT;   204 AA.
AC   O08826;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Axonemal dynein heavy chain (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NMR1; TISSUE=Testis;
RX   MEDLINE=98038992; PubMed=9373155;
RA   Neesen J., Koehle M., Kirschne R., Steinlein C., Kreutzberger J.,
RA   Engel W., Schmid M.;
RT   "Identification of dynein heavy chain genes expressed in human and
RT   mouse testis: chromosomal localization of an axonemal dynein gene.";
RL   Gene 200:193-202(1997).
DR   EMBL; Z83813; CAB06067.1; -.
FT   NON_TER       1      1
FT   NON_TER     204    204
SQ   SEQUENCE   204 AA;  22449 MW;  9AFFDA82E4603746 CRC64;

  Query Match             100.0%;  Score 28;  DB 11;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        199 DYGMS 203


RESULT 5
O08823
ID   O08823      PRELIMINARY;      PRT;   204 AA.
AC   O08823;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Axonemal dynein heavy chain (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NMR1; TISSUE=Testis;
RX   MEDLINE=98038992; PubMed=9373155;
RA   Neesen J., Koehle M., Kirschne R., Steinlein C., Kreutzberger J.,
RA   Engel W., Schmid M.;
RT   "Identification of dynein heavy chain genes expressed in human and
RT   mouse testis: chromosomal localization of an axonemal dynein gene.";
RL   Gene 200:193-202(1997).
DR   EMBL; Z83810; CAB06064.1; -.
FT   NON_TER       1      1
FT   NON_TER     204    204
SQ   SEQUENCE   204 AA;  22198 MW;  72756AAFB52D0B84 CRC64;

  Query Match             100.0%;  Score 28;  DB 11;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        199 DYGMS 203


RESULT 6
Q8ZJU0
ID   Q8ZJU0      PRELIMINARY;      PRT;   227 AA.
AC   Q8ZJU0;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative fimbrial chaparone protein.
GN   STHA OR STM4594.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
```

```
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
DR   EMBL; AE008916; AAL23409.1; -.
DR   GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR   GO; GO:0003754; F:chaperone activity; IEA.
DR   GO; GO:0007047; P:cell wall organization and biogenesis; IEA.
DR   InterPro; IPR008962; PapD-like.
DR   InterPro; IPR001829; Pili_chaperone.
DR   Pfam; PF00345; pili_assembly; 1.
DR   Pfam; PF02753; pili_assembly_C; 1.
DR   PRINTS; PR00969; CHAPERONPILI.
DR   ProDom; PD001447; Pili_chaperone; 1.
DR   PROSITE; PS00635; PILI_CHAPERONE; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   227 AA;  24743 MW;  EC0D30D9487CD7DF CRC64;

  Query Match             100.0%;  Score 28;  DB 16;  Length 227;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        215 DYGMS 219


RESULT 7
Q8Z0S8
ID   Q8Z0S8      PRELIMINARY;      PRT;   227 AA.
AC   Q8Z0S8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative fimbrial chaperone protein.
GN   STHB OR STY4943 OR T4633.
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Ty2 / ATCC 700931;
RX   MEDLINE=22531367; PubMed=12644504;
RA   Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA   Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT   "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT   and CT18.";
RL   J. Bacteriol. 185:2330-2337(2003).
DR   EMBL; AL627284; CAD03425.1; -.
DR   EMBL; AE016849; AAO72062.1; -.
DR   GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR   GO; GO:0003754; F:chaperone activity; IEA.
DR   GO; GO:0007047; P:cell wall organization and biogenesis; IEA.
DR   InterPro; IPR008962; PapD-like.
DR   InterPro; IPR001829; Pili_chaperone.
DR   Pfam; PF00345; pili_assembly; 1.
DR   Pfam; PF02753; pili_assembly_C; 1.
DR   PRINTS; PR00969; CHAPERONPILI.
DR   ProDom; PD001447; Pili_chaperone; 1.
DR   PROSITE; PS00635; PILI_CHAPERONE; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   227 AA;  24829 MW;  6E8F92D7B5421AD1 CRC64;

  Query Match             100.0%;  Score 28;  DB 16;  Length 227;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        215 DYGMS 219


RESULT 8
Q8E110
ID   Q8E110      PRELIMINARY;      PRT;   249 AA.
AC   Q8E110;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Prophage LambdaSa1, antirepressor, putative.
GN   SAG0555.
OS   Streptococcus agalactiae (serotype V).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=216466;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2603 V/R / Serotype V;
RX   MEDLINE=22222988; PubMed=12200547;
RA   Tettelin H., Masignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA   Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA   Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA   DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA   Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA   Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA   Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,
RA   Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA   Fraser C.M.;
RT   "Complete genome sequence and comparative genomic analysis of an
RT   emerging human pathogen, serotype V Streptococcus agalactiae.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR   EMBL; AE014217; AAM99456.1; -.
DR   TIGR; SAG0555; -.
KW   Complete proteome.
SQ   SEQUENCE   249 AA;  29176 MW;  F6537740D2C89258 CRC64;

  Query Match             100.0%;  Score 28;  DB 16;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        163 DYGMS 167


RESULT 9
Q9SKM6
ID   Q9SKM6      PRELIMINARY;      PRT;   268 AA.
AC   Q9SKM6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   En/Spm-like transposon protein.
GN   AT2G28440.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC006283; AAD20682.1; -.
DR   PIR; H84684; H84684.
SQ   SEQUENCE   268 AA;  27741 MW;  0584FDE4622FED0E CRC64;

  Query Match             100.0%;  Score 28;  DB 10;  Length 268;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        218 DYGMS 222


RESULT 10
O81788
ID   O81788      PRELIMINARY;      PRT;   268 AA.
AC   O81788;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Homeodomain-like protein (AT4G35550/F8D20_60).
GN   F8D20.60 OR AT4G35550.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Koetter P., Hempel S., Entian K.-D., Hoheisel J., Jesse T.,
RA   Heijnen L., Vos P., Mewes H.W., Mayer K.F.X., Schueller C., Bevan M.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA   Banh J., Bowser L., Carninci P., Dale J.M., Gibson H.A.,
RA   Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J.,
RA   Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA   Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA   Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
RA   Davis R.W., Theologis A., Ecker J.R.;
RT   "Arabidopsis cDNA clones.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA   Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA   Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA   Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA   Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA   Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA   Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
RA   Ecker J.R.;
RT   "Arabidopsis ORF clones.";
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL031135; CAA20025.1; -.
DR   EMBL; AL161587; CAB80271.1; -.
DR   EMBL; AY048268; AAK82530.1; -.
DR   EMBL; BT000538; AAN18107.1; -.
DR   PIR; T04660; T04660.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
SQ   SEQUENCE   268 AA;  29673 MW;  D57FCC13AB1A93DA CRC64;

  Query Match             100.0%;  Score 28;  DB 10;  Length 268;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        263 DYGMS 267


RESULT 11
Q8PFN0
ID   Q8PFN0      PRELIMINARY;      PRT;   293 AA.
AC   Q8PFN0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein XAC3950.
GN   XAC3950.
OS   Xanthomonas axonopodis (pv. citri).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=92829;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=306 / ATCC 13902 / XV 101;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
```

```
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
DR   EMBL; AE012044; AAM38787.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   293 AA;  33115 MW;  8CAC3C831EA0B33F CRC64;

  Query Match             100.0%;  Score 28;  DB 16;  Length 293;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        234 DYGMS 238


RESULT 12
Q9C840
ID   Q9C840      PRELIMINARY;      PRT;   300 AA.
AC   Q9C840;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   T8E24.2.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016720; PubMed=11130713;
RA   Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M.,
RA   Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA   Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA   De Simone V., Choisne N., Artiguenave F., Robert C., Brottier P.,
RA   Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA   Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA   Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA   Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA   Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA   Conrad A., Hornischer K., Kauer G., Loehnert T.-H., Nordsiek G.,
RA   Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
RA   Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA   Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA   de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA   Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA   Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA   Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA   Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA   Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA   Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA   Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA   Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA   Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA   Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA   Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT   "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:820-822(2000).
DR   EMBL; AC036106; AAG51001.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   300 AA;  33773 MW;  C05A46267AFCE3F4 CRC64;

  Query Match             100.0%;  Score 28;  DB 10;  Length 300;
  Best Local Similarity   100.0%;  Pred. No. 2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         32 DYGMS 36


RESULT 13
Q9M7X8
ID   Q9M7X8      PRELIMINARY;      PRT;   302 AA.
AC   Q9M7X8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   F3E22.15 protein.
GN   F3E22.15.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA   Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA   Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT   "Arabidopsis thaliana chromosome III BAC F3E22 genomic sequence.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC023912; AAF63827.1; -.
SQ   SEQUENCE   302 AA;  34042 MW;  C7C97F75F2825AE9 CRC64;

  Query Match             100.0%;  Score 28;  DB 10;  Length 302;
  Best Local Similarity   100.0%;  Pred. No. 2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         32 DYGMS 36


RESULT 14
Q80YY6
ID   Q80YY6      PRELIMINARY;      PRT;   333 AA.
AC   Q80YY6;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   BM289L21.4.1 (Novel protein, variant 1) (Fragment).
GN   BM289L21.4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   North P., Leaves N., Greystrong J., Coppola M., Manjunath S.,
RA   Russell E., Smith M., Strachan G., Tofts C., Boal E., Cobley V.,
RA   Hunter G., Kimberley C., Thomas D., Cave-Berry L., Weston P.,
RA   Botcherby M.R.M.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX247953; CAD83027.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   333 AA;  38205 MW;  42A7BDF256DA90C2 CRC64;

  Query Match             100.0%;  Score 28;  DB 11;  Length 333;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db        272 DYGMS 276
```

```
RESULT 15
Q9YDP1
ID   Q9YDP1      PRELIMINARY;      PRT;   340 AA.
AC   Q9YDP1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein APE0874.
GN   APE0874.
OS   Aeropyrum pernix.
OC   Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC   Desulfurococcaceae; Aeropyrum.
OX   NCBI_TaxID=56636;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K1;
RX   MEDLINE=99310339; PubMed=10382966;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA   Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA   Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA   Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA   Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA   Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT   "Complete genome sequence of an aerobic hyper-thermophilic
RT   crenarchaeon, Aeropyrum pernix K1.";
RL   DNA Res. 6:83-101(1999).
DR   EMBL; AP000060; BAA79856.1; -.
DR   PIR; H72681; H72681.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   340 AA;  39481 MW;  369F8FA81963C3EF CRC64;

  Query Match             100.0%;  Score 28;  DB 17;  Length 340;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DYGMS 5
              |||||
Db         82 DYGMS 86


Search completed: September 18, 2004, 03:46:26
Job time : 29.0714 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 02:28:24 ; Search time 120.214 Seconds
(without alignments)
39.956 Million cell updates/sec

Title: US-10-029-926B-115
Perfect score: 96
Sequence: 1 GINWNGGSTGYADSVKG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	96	100.0	17	3	AAY95192	Aay95192 Anti-plat
2	96	100.0	17	5	ABG78240	Abg78240 Human Fv
3	96	100.0	17	5	ABG91931	Abg91931 Human ant
4	96	100.0	98	3	AAB40073	Aab40073 Anti-hIL1
5	96	100.0	98	5	ABG78186	Abg78186 Human Fv
6	96	100.0	98	5	ABG91877	Abg91877 Human ant
7	96	100.0	98	6	ABO27087	Abo27087 Human ger
8	96	100.0	113	3	AAY95177	Aay95177 Anti-plat
9	96	100.0	113	3	AAY95178	Aay95178 Anti-plat
10	96	100.0	115	3	AAY95189	Aay95189 Anti-plat
11	96	100.0	115	3	AAY95190	Aay95190 Anti-plat
12	96	100.0	116	2	AAW19880	Aaw19880 CEA-speci
13	96	100.0	117	2	AAR66312	Aar66312 Human imm
14	96	100.0	118	4	AAU02560	Aau02560 Anti-adip
15	96	100.0	122	3	AAY96065	Aay96065 Human ant
16	96	100.0	207	5	AAU98019	Aau98019 Human ace
17	96	100.0	238	3	AAY95198	Aay95198 Anti-plat
18	96	100.0	239	5	ABP46004	Abp46004 Human BLy
19	96	100.0	239	5	ABP46027	Abp46027 Human BLy
20	96	100.0	239	5	ABP44926	Abp44926 Human BLy
21	96	100.0	244	6	AAO31136	Aao31136 Human CM0
22	96	100.0	246	5	ABG78329	Abg78329 Human Fv
23	96	100.0	246	5	ABG92026	Abg92026 Antibody
24	96	100.0	256	5	ABG78334	Abg78334 Human Fv
25	96	100.0	256	5	ABG92025	Abg92025 Antibody
26	96	100.0	266	5	ABG92020	Abg92020 Human ant
27	96	100.0	277	5	ABG78150	Abg78150 Human Fv
28	96	100.0	277	5	ABG78328	Abg78328 Human Fv
29	96	100.0	277	5	ABG92019	Abg92019 Human ant
30	96	100.0	277	5	ABG91841	Abg91841 Human ant
31	96	100.0	309	2	AAW83322	Aaw83322 Single ch
32	96	100.0	309	5	ABB09603	Abb09603 Amino aci
33	96	100.0	309	6	ABG74384	Abg74384 Single ch
34	96	100.0	464	5	ABG78151	Abg78151 Human Fv
35	96	100.0	464	5	ABG91842	Abg91842 Human ant
36	93	96.9	126	4	AAU02603	Aau02603 Anti-adip
37	91	94.8	17	3	AAB52196	Aab52196 Human ant
38	91	94.8	128	3	AAB52199	Aab52199 Human ant
39	90	93.8	464	5	ABG92021	Abg92021 Antibody
40	79	82.3	98	3	AAB40072	Aab40072 Anti-hIL1
41	79	82.3	98	5	ABG78185	Abg78185 Human Fv
42	79	82.3	98	5	ABG91876	Abg91876 Human ant
43	79	82.3	99	3	AAY50960	Aay50960 Human FVI
44	79	82.3	99	6	ABO27083	Abo27083 Human ger
45	79	82.3	114	4	AAU02592	Aau02592 Anti-adip

ALIGNMENTS

RESULT 1
AAY95192
ID AAY95192 standard; peptide; 17 AA.
XX
AC AAY95192;
XX
DT 29-AUG-2000 (first entry)
XX
DE Anti-platelet glycoprotein Ib human HIb-1 VH CDR2.
XX
KW Variable heavy chain; single chain antibody; scFv; human; HIb-1;
KW glycoprotein Ib alpha; platelet; aggregation; antiaggregant;
KW antithrombotic; thrombus; therapy; diagnostic; CDR2;
KW complementarity determining region.
XX
OS Homo sapiens.
XX
PN WO200026667-A1.
XX
PD 11-MAY-2000.
XX
PF 29-OCT-1999; 99WO-US025495.
XX
PR 30-OCT-1998; 98US-0106275P.
XX
PA (MILL/) MILLER J L.
XX
PI Miller JL;
XX
DR WPI; 2000-365744/31.
XX
PT Isolated nucleic acid molecule encoding anti-human platelet glycoprotein
PT Ib alpha molecule useful for producing antibodies which inhibit platelet
PT aggregation.
XX
PS Claim 15; Fig 5; 89pp; English.
XX
CC The present sequence is that of complementarity determining region 2
CC (CDR2) of the heavy chain variable region (VH) of human single chain
CC antibody (scFv) HIb-1 (see AAY95198), which is directed against platelet
CC glycoprotein Ib (GPIb). The HIb series of scFv was isolated from a human
CC synthetic VH and VL scFv library on the basis of their binding to
CC platelet GPIb. Whether displayed as surface proteins on a phagemid or
CC secreted as free scFv by Escherichia coli, the HIb scFv clones are
CC capable of inhibiting von Willebrand factor-dependent aggregation of
CC platelets. The scFv are composed of native human protein sequences and
CC are therefore attractive potential reagents for therapeutic purposes.
CC They provide a new class of antithrombotic agents, useful for the

```
CC   prevention of platelet-dependent thrombi in diseased arteries, bypass
CC   grafts, dialysis etc., and can also be used as diagnostic reagents.
CC   Methods of inhibiting aggregation of platelets, of binding human platelet
CC   GPIb alpha and of selecting a VH or VL region of an antibody that
CC   inhibits platelet aggregation are claimed. Fragments of the scFv VH or VL
CC   chain, including CDR fragments, are also claimed
XX
SQ   Sequence 17 AA;

  Query Match             100.0%;  Score 96;  DB 3;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-08;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db          1 GINWNGGSTGYADSVKG 17


RESULT 2
ABG78240
ID   ABG78240 standard; peptide; 17 AA.
XX
AC   ABG78240;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Human Fv molecule hypervariable region related peptide #115.
XX
KW   Human; Fv molecule; hypervariable region; single chain Fv; cytostatic;
KW   disulfide Fv; dsFv; scFv; cancer; carcinoma; sarcoma; leukaemia; adenoma;
KW   lymphoma; myeloma; blastoma; seminoma; melanoma; acute myeloid leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200259264-A2.
XX
PD   01-AUG-2002.
XX
PF   31-DEC-2001; 2001WO-US049440.
XX
PR   29-DEC-2000; 2000US-00751181.
XX
PA   (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI   Hagai Y,  Lazarovits J,  Guy R,  Lipschitz O,  Szanton E,  Levanon A;
PI   Plaksin D,  Peretz T;
XX
DR   WPI; 2002-619166/66.
XX
PT   Novel peptide/polypeptide for cancer therapy has Fv molecule, construct
PT   or fragment, or construct of fragment with enhanced binding
PT   characteristics so as to selectively bind target cell in favor of other
PT   cells.
XX
PS   Claim 20; Page 208-209; 232pp; English.
XX
CC   The invention relates to a peptide or polypeptide comprising an Fv
CC   molecule, a construct or fragments or a construct of a fragment with
CC   enhanced binding characteristics which selectively and/or specifically
CC   binds to a target cell in favour of other cells, where binding is
CC   primarily determined by a first hypervariable region and Fv is a single
CC   chain Fv (scFv) or a disulfide Fv (dsFv). The peptide, optionally in
CC   association with or attached, coupled, combined, linked or fused to a
CC   pharmaceutical agent, is useful in the manufacture of a medicament, where
CC   the medicament has activity against a diseased cell, preferably a cancer
CC   cell (selected from carcinoma, sarcoma, leukaemia, adenoma, lymphoma,
CC   myeloma, blastoma, seminoma, and melanoma, where the leukaemia cell is an
CC   acute myeloid leukaemia cell). The peptide is also useful for preparing a
CC   composition for use in inhibiting the growth of a diseased or cancer
CC   cell. This sequence represents a human Fv molecule hypervariable region
CC   related peptide of the invention
XX
SQ   Sequence 17 AA;

  Query Match             100.0%;  Score 96;  DB 5;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-08;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db          1 GINWNGGSTGYADSVKG 17


RESULT 3
ABG91931
ID   ABG91931 standard; peptide; 17 AA.
XX
AC   ABG91931;
XX
DT   04-DEC-2002  (first entry)
XX
DE   Human antibody fragment #115.
XX
KW   Human; antibody; epitope; cancer; tumour; cell rolling; inflammation;
KW   metastasis; hypervariable region; autoimmune disease; thrombosis;
KW   restenosis; leukaemia; inflammatory disease; cardiovascular disease;
KW   myocardial infarction; retinopathic disease; abnormal platelet function;
KW   sulphated tyrosine-dependent protein-protein interaction.
XX
OS   Homo sapiens.
XX
PN   WO200253700-A2.
XX
PD   11-JUL-2002.
XX
PF   31-DEC-2001; 2001WO-US049442.
XX
PR   29-DEC-2000; 2000US-00751181.
PR   29-DEC-2000; 2000US-0258948P.
XX
PA   (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI   Lazarovits J,  Hagai Y,  Plaksin D,  Vogel T,  Nimrod A,  Mar-Haim H;
PI   Szanthon E,  Richter T,  Amit B,  Kooperman L,  Peretz T,  Levanon A;
XX
DR   WPI; 2002-674776/72.
XX
PT   Novel isolated epitope present on cancer cells and important in
PT   physiological phenomena such as cell rolling, metastasis and
PT   inflammation, for treating autoimmune, inflammatory or cardiovascular
PT   diseases, and cancer.
XX
PS   Claim 25; Page 284-285; 0pp; English.
XX
CC   The invention relates to an isolated epitope present on cancer cells and
CC   important in physiological phenomena such as cell rolling, metastasis and
CC   inflammation, where the epitope is capable of being bound by an antibody,
CC   its antigen-binding fragment or its complex comprising at least one
CC   antibody or its binding fragment having a first hypervariable region. The
CC   epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
CC   disease, thrombosis, restenosis, metastasis, growth and/or replication of
CC   tumour or leukaemia cells, increase in number of tumour or leukaemia
CC   cells in a patient, cell-cell, cell-matrix, platelet-matrix, platelet-
CC   platelet and/or cell-platelet adhesion or aggregation, for increasing
CC   mortality of tumour or leukaemia cells, for increasing the susceptibility
CC   of diseased cells to damage by anti-disease, anti-cancer or anti-
CC   leukaemia agents, or for decreasing the number of tumour or leukaemia
CC   cells in a patient, or in the manufacture of a medicament for the above
CC   mentioned purposes. The epitopes are useful for diagnosing and treating
CC   diseases such as cancer, leukaemia, autoimmune diseases, inflammatory
CC   diseases, cardiovascular diseases such as myocardial infarction,
CC   retinopathic diseases and other diseases mediated by abnormal platelet
CC   function and diseases caused by sulphated tyrosine-dependent protein-
CC   protein interactions. This sequence represents a human antibody fragment
CC   of the invention
XX
```

SQ Sequence 17 AA;

Query Match 100.0%; Score 96; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 9.5e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 1 GINWNGGSTGYADSVKG 17

RESULT 4
AAB40073
ID AAB40073 standard; protein; 98 AA.
XX
AC AAB40073;
XX
DT 05-FEB-2001 (first entry)
XX
DE Anti-hIL12 antibody H chain V region amino acid sequence SEQ ID 599.
XX
KW Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
KW complementarity determining region; CDR; antirheumatic; antiarthritic;
KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiant;
KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200056772-A1.
XX
PD 28-SEP-2000.
XX
PF 24-MAR-2000; 2000WO-US007946.
XX
PR 25-MAR-1999; 99US-0126603P.
XX
PA (BADI) BASF AG.
PA (GEMY) GENETICS INST INC.
XX
PI Salfeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
PI Kaymakcalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
PI Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR;
PI Derbyshire EJ, Carmen S, Smith S, Holtet TL, Du Fou SL;
XX
DR WPI; 2000-638250/61.
XX
PT New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT disease and multiple sclerosis.
XX
PS Claim 75; Page 121; 377pp; English.
XX
CC This invention relates to a new human antibody specific for human
CC interleukin-12 (IL-12). The invention also includes antigen binding
CC portions that bind to IL-12. Sequences AAB39485-B39516 represent human
CC anti-IL-12 antibody heavy and light chain complementarity determining
CC region (CDR) amino acid sequences, and also includes variable region
CC amino acid sequences. Other variable region amino acid sequences are
CC given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771
CC represent anti-IL-12 CDR3 related amino acid sequences, AAB39772-B40063
CC represent other CDR sequences. Light chain CDR3 consensus sequences are
CC given in AAB40064-B40067. Primers used in the identification and
CC construction of the antibodies of the invention are given in AAC61062-
CC C61071. The antibody of the invention is a neutralising antibody and has
CC antirheumatic; antiarthritic; antisclerotic; antiinflammatory;
CC neuroprotective; antipsoriatic; antiasthmatic; cardiant; antiparasitic;
CC antibacterial and immunosuppressive activity. The antibodies or antigen-
CC binding fragments are useful in the treatment of disorders associated
CC with detrimental release of human IL-12, especially Crohn's disease,
CC multiple sclerosis and rheumatoid arthritis. They can also be used in the
CC manufacture of a pharmaceutical composition to treat human IL-12
CC disorders
XX
SQ Sequence 98 AA;

Query Match 100.0%; Score 96; DB 3; Length 98;
Best Local Similarity 100.0%; Pred. No. 6.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 50 GINWNGGSTGYADSVKG 66

RESULT 5
ABG78186
ID ABG78186 standard; protein; 98 AA.
XX
AC ABG78186;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human Fv molecule hypervariable region related peptide #61.
XX
KW Human; Fv molecule; hypervariable region; single chain Fv; cytostatic;
KW disulfide Fv; dsFv; scFv; cancer; carcinoma; sarcoma; leukaemia; adenoma;
KW lymphoma; myeloma; blastoma; seminoma; melanoma; acute myeloid leukaemia.
XX
OS Homo sapiens.
XX
PN WO200259264-A2.
XX
PD 01-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049440.
XX
PR 29-DEC-2000; 2000US-00751181.
XX
PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI Hagai Y, Lazarovits J, Guy R, Lipschitz O, Szanton E, Levanon A;
PI Plaksin D, Peretz T;
XX
DR WPI; 2002-619166/66.
XX
PT Novel peptide/polypeptide for cancer therapy has Fv molecule, construct
PT or fragment, or construct of fragment with enhanced binding
PT characteristics so as to selectively bind target cell in favor of other
PT cells.
XX
PS Claim 13; Page 177-178; 232pp; English.
XX
CC The invention relates to a peptide or polypeptide comprising an Fv
CC molecule, a construct or fragments or a construct of a fragment with
CC enhanced binding characteristics which selectively and/or specifically
CC binds to a target cell in favour of other cells, where binding is
CC primarily determined by a first hypervariable region and Fv is a single
CC chain Fv (scFv) or a disulfide Fv (dsFv). The peptide, optionally in
CC association with or attached, coupled, combined, linked or fused to a
CC pharmaceutical agent, is useful in the manufacture of a medicament, where
CC the medicament has activity against a diseased cell, preferably a cancer
CC cell (selected from carcinoma, sarcoma, leukaemia, adenoma, lymphoma,
CC myeloma, blastoma, seminoma, and melanoma, where the leukaemia cell is an
CC acute myeloid leukaemia cell). The peptide is also useful for preparing a
CC composition for use in inhibiting the growth of a diseased or cancer
CC cell. This sequence represents a human Fv molecule hypervariable region
CC related peptide of the invention
XX
SQ Sequence 98 AA;

Query Match 100.0%; Score 96; DB 5; Length 98;
Best Local Similarity 100.0%; Pred. No. 6.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17

```
                  ||||||||||||||||||
Db          50 GINWNGGSTGYADSVKG 66


RESULT 6
ABG91877
ID   ABG91877 standard; protein; 98 AA.
XX
AC   ABG91877;
XX
DT   04-DEC-2002  (first entry)
XX
DE   Human antibody fragment #61.
XX
KW   Human; antibody; epitope; cancer; tumour; cell rolling; inflammation;
KW   metastasis; hypervariable region; autoimmune disease; thrombosis;
KW   restenosis; leukaemia; inflammatory disease; cardiovascular disease;
KW   myocardial infarction; retinopathic disease; abnormal platelet function;
KW   sulphated tyrosine-dependent protein-protein interaction.
XX
OS   Homo sapiens.
XX
PN   WO200253700-A2.
XX
PD   11-JUL-2002.
XX
PF   31-DEC-2001; 2001WO-US049442.
XX
PR   29-DEC-2000; 2000US-00751181.
PR   29-DEC-2000; 2000US-0258948P.
XX
PA   (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI   Lazarovits J,  Hagai Y,  Plaksin D,  Vogel T,  Nimrod A,  Mar-Haim H;
PI   Szanthon E,  Richter T,  Amit B,  Kooperman L,  Peretz T,  Levanon A;
XX
DR   WPI; 2002-674776/72.
XX
PT   Novel isolated epitope present on cancer cells and important in
PT   physiological phenomena such as cell rolling, metastasis and
PT   inflammation, for treating autoimmune, inflammatory or cardiovascular
PT   diseases, and cancer.
XX
PS   Disclosure; Page 255; 0pp; English.
XX
CC   The invention relates to an isolated epitope present on cancer cells and
CC   important in physiological phenomena such as cell rolling, metastasis and
CC   inflammation, where the epitope is capable of being bound by an antibody,
CC   its antigen-binding fragment or its complex comprising at least one
CC   antibody or its binding fragment having a first hypervariable region. The
CC   epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
CC   disease, thrombosis, restenosis, metastasis, growth and/or replication of
CC   tumour or leukaemia cells, increase in number of tumour or leukaemia
CC   cells in a patient, cell-cell, cell-matrix, platelet-matrix, platelet-
CC   platelet and/or cell-platelet adhesion or aggregation, for increasing
CC   mortality of tumour or leukaemia cells, for increasing the susceptibility
CC   of diseased cells to damage by anti-disease, anti-cancer or anti-
CC   leukaemia agents, or for decreasing the number of tumour or leukaemia
CC   cells in a patient, or in the manufacture of a medicament for the above
CC   mentioned purposes. The epitopes are useful for diagnosing and treating
CC   diseases such as cancer, leukaemia, autoimmune diseases, inflammatory
CC   diseases, cardiovascular diseases such as myocardial infarction,
CC   retinopathic diseases and other diseases mediated by abnormal platelet
CC   function and diseases caused by sulphated tyrosine-dependent protein-
CC   protein interactions. This sequence represents a human antibody fragment
CC   of the invention
XX
SQ   Sequence 98 AA;

  Query Match            100.0%;  Score 96;  DB 5;  Length 98;
  Best Local Similarity  100.0%;  Pred. No. 6.7e-07;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
                ||||||||||||||||||
Db          50 GINWNGGSTGYADSVKG 66


RESULT 7
ABO27087
ID   ABO27087 standard; protein; 98 AA.
XX
AC   ABO27087;
XX
DT   10-SEP-2003  (first entry)
XX
DE   Human germline heavy chain variable region gene segment #20.
XX
KW   Human; heavy chain variable region; VH; humanised antibody;
KW   chimeric antibody; complementarity determining region; CDR;
KW   canonical CDR structure type.
XX
OS   Homo sapiens.
XX
PN   US2003039649-A1.
XX
PD   27-FEB-2003.
XX
PF   12-JUL-2002; 2002US-00194975.
XX
PR   12-JUL-2001; 2001US-0305111P.
XX
PA   (FOOT/) FOOTE J.
XX
PI   Foote J;
XX
DR   WPI; 2003-492151/46.
XX
PT   Making humanized antibody for converting antibody, by making chimeric
PT   antibodies containing complementarity determining region from non-human
PT   antibody and appropriate framework sequences of human antibodies.
XX
PS   Example 1; Fig 1; 31pp; English.
XX
CC   The invention describes a method of making a humanised antibody,
CC   comprising making chimeric antibodies containing a complementarity
CC   determining region (CDR) from a non-human antibody and appropriate
CC   framework sequences (I) of human antibodies. (I) is selected by using
CC   canonical CDR structure types of non-human antibody in comparison to
CC   germline canonical CDR structure types of human antibodies as the basis
CC   for selection, for humanisation. The method is useful for making a
CC   humanised antibody or a converted antibody. The method is applicable for
CC   converting a subject antibody sequence of any subject species to a less
CC   immunogenic form suitable for use in an object species. The method is
CC   reliable for identifying suitable human framework sequences to support
CC   non-human CDR regions and to provide humanised antibodies that retain
CC   high antigen binding with low immunogenicity in humans, without the need
CC   for direct comparison of framework sequences, without the need for
CC   determining critically important amino acid residues in the framework,
CC   and without the need for multiple iteration and construction to obtain
CC   humanised antibodies with suitable therapeutic properties. The antibody
CC   has high affinity and low immunogenicity without need for comparing
CC   framework sequences between non-human and human antibodies. This sequence
CC   represents a human heavy chain variable region gene segment used in the
CC   creation of humanised antibodies
XX
SQ   Sequence 98 AA;

  Query Match            100.0%;  Score 96;  DB 6;  Length 98;
  Best Local Similarity  100.0%;  Pred. No. 6.7e-07;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
                ||||||||||||||||||
Db          50 GINWNGGSTGYADSVKG 66
```

```
RESULT 8
AAY95177
ID   AAY95177 standard; protein; 113 AA.
XX
AC   AAY95177;
XX
DT   29-AUG-2000  (first entry)
XX
DE   Anti-platelet glycoprotein Ib human HIb-1 VH.
XX
KW   Variable heavy chain; single chain antibody; scFv; human; HIb-1;
KW   glycoprotein Ib alpha; platelet; aggregation; antiaggregant;
KW   antithrombotic; thrombus; therapy; diagnostic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .30
FT                   /note= "framework region 1"
FT   Region          31. .34
FT                   /note= "complementarity determining region 1"
FT   Region          35. .49
FT                   /note= "framework region 2"
FT   Region          50. .66
FT                   /note= "complementarity determining region 2"
FT   Region          67. .98
FT                   /note= "framework region 3"
FT   Region          99. .104
FT                   /note= "complementarity determining region 3"
FT   Region          105. .113
FT                   /note= "framework region 4"
XX
PN   WO200026667-A1.
XX
PD   11-MAY-2000.
XX
PF   29-OCT-1999;   99WO-US025495.
XX
PR   30-OCT-1998;   98US-0106275P.
XX
PA   (MILL/) MILLER J L.
XX
PI   Miller JL;
XX
DR   WPI; 2000-365744/31.
DR   N-PSDB; AAA27658.
XX
PT   Isolated nucleic acid molecule encoding anti-human platelet glycoprotein
PT   Ib alpha molecule useful for producing antibodies which inhibit platelet
PT   aggregation.
XX
PS   Claim 10; Page 71; 89pp; English.
XX
CC   The present sequence is that of the heavy chain variable region (VH) of
CC   human single chain antibody (scFv) HIb-1 (see AAY95198), which is
CC   directed against platelet glycoprotein Ib (GPIb). The HIb series of scFv
CC   was isolated from a human synthetic VH and VL scFv library by 3 rounds of
CC   phagemid selection against transfected CHO cells expressing the GPIb
CC   alpha component of the GPIb/IX/V complex on their surface, followed by a
CC   4th round of selection against washed human platelets, and 2 final rounds
CC   in which attempts were made to displace scFv from washed platelets by
CC   flooding with murine monoclonal antibody or mimotope peptide (see
CC   AAY95229). Whether displayed as surface proteins on a phagemid or
CC   secreted as free scFv by Escherichia coli, the HIb scFv clones are
CC   capable of inhibiting von Willebrand factor-dependent aggregation of
CC   platelets. The scFv are composed of native human protein sequences and
CC   are therefore attractive potential reagents for therapeutic purposes.
CC   They provide a new class of antithrombotic agents, useful for the
CC   prevention of platelet-dependent thrombi in diseased arteries, bypass
CC   grafts, dialysis etc., and can also be used as diagnostic reagents.
CC   Methods of inhibiting aggregation of platelets, of binding human platelet
CC   GPIb alpha and of selecting a VH or VL region of an antibody that
CC   inhibits platelet aggregation are claimed
XX
SQ   Sequence 113 AA;

  Query Match             100.0%;  Score 96;  DB 3;  Length 113;
  Best Local Similarity   100.0%;  Pred. No. 7.8e-07;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         50 GINWNGGSTGYADSVKG 66


RESULT 9
AAY95178
ID   AAY95178 standard; protein; 113 AA.
XX
AC   AAY95178;
XX
DT   29-AUG-2000  (first entry)
XX
DE   Anti-platelet glycoprotein Ib human HIb-1 VH.
XX
KW   Variable heavy chain; single chain antibody; scFv; human; HIb-1;
KW   glycoprotein Ib alpha; platelet; aggregation; antiaggregant;
KW   antithrombotic; thrombus; therapy; diagnostic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .30
FT                   /note= "framework region 1"
FT   Region          31. .34
FT                   /note= "complementarity determining region 1"
FT   Region          35. .49
FT                   /note= "framework region 2"
FT   Region          50. .66
FT                   /note= "complementarity determining region 2"
FT   Region          67. .98
FT                   /note= "framework region 3"
FT   Region          99. .104
FT                   /note= "complementarity determining region 3"
FT   Region          105. .113
FT                   /note= "framework region 4"
XX
PN   WO200026667-A1.
XX
PD   11-MAY-2000.
XX
PF   29-OCT-1999;   99WO-US025495.
XX
PR   30-OCT-1998;   98US-0106275P.
XX
PA   (MILL/) MILLER J L.
XX
PI   Miller JL;
XX
DR   WPI; 2000-365744/31.
XX
PT   Isolated nucleic acid molecule encoding anti-human platelet glycoprotein
PT   Ib alpha molecule useful for producing antibodies which inhibit platelet
PT   aggregation.
XX
PS   Claim 10; Page; 89pp; English.
XX
CC   The present sequence is that of the heavy chain variable region (VH) of
CC   human single chain antibody (scFv) HIb-1 (see AAY95198), which is
CC   directed against platelet glycoprotein Ib (GPIb). The HIb series of scFv
CC   was isolated from a human synthetic VH and VL scFv library by 3 rounds of
CC   phagemid selection against transfected CHO cells expressing the GPIb
CC   alpha component of the GPIb/IX/V complex on their surface, followed by a
CC   4th round of selection against washed human platelets, and 2 final rounds
CC   in which attempts were made to displace scFv from washed platelets by
```

```
CC   flooding with murine monoclonal antibody or mimotope peptide (see
CC   AAY95229). Whether displayed as surface proteins on a phagemid or
CC   secreted as free scFv by Escherichia coli, the HIb scFv clones are
CC   capable of inhibiting von Willebrand factor-dependent aggregation of
CC   platelets. The scFv are composed of native human protein sequences and
CC   are therefore attractive potential reagents for therapeutic purposes.
CC   They provide a new class of antithrombotic agents, useful for the
CC   prevention of platelet-dependent thrombi in diseased arteries, bypass
CC   grafts, dialysis etc., and can also be used as diagnostic reagents.
CC   Methods of inhibiting aggregation of platelets, of binding human platelet
CC   GPIb alpha and of selecting a VH or VL region of an antibody that
CC   inhibits platelet aggregation are claimed. Note: The present sequence is
CC   not shown in the specification but is derived from the HIb-1 VH sequence
CC   given on page 71 (see AAY95177)
XX
SQ   Sequence 113 AA;

  Query Match             100.0%;  Score 96;  DB 3;  Length 113;
  Best Local Similarity   100.0%;  Pred. No. 7.8e-07;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         50 GINWNGGSTGYADSVKG 66


RESULT 10
AAY95189
ID   AAY95189 standard; protein; 115 AA.
XX
AC   AAY95189;
XX
DT   29-AUG-2000  (first entry)
XX
DE   Anti-platelet glycoprotein Ib human HIb-1 VH.
XX
KW   Variable heavy chain; single chain antibody; scFv; human; HIb-1;
KW   glycoprotein Ib alpha; platelet; aggregation; antiaggregant;
KW   antithrombotic; thrombus; therapy; diagnostic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .30
FT                   /note= "framework region 1"
FT   Region          31. .34
FT                   /note= "complementarity determining region 1"
FT   Region          35. .49
FT                   /note= "framework region 2"
FT   Region          50. .66
FT                   /note= "complementarity determining region 2"
FT   Region          67. .98
FT                   /note= "framework region 3"
FT   Region          99. .104
FT                   /note= "complementarity determining region 3"
FT   Region          105. .115
FT                   /note= "framework region 4"
XX
PN   WO200026667-A1.
XX
PD   11-MAY-2000.
XX
PF   29-OCT-1999;   99WO-US025495.
XX
PR   30-OCT-1998;   98US-0106275P.
XX
PA   (MILL/) MILLER J L.
XX
PI   Miller JL;
XX
DR   WPI; 2000-365744/31.
XX
PT   Isolated nucleic acid molecule encoding anti-human platelet glycoprotein
PT   Ib alpha molecule useful for producing antibodies which inhibit platelet
PT   aggregation.
XX
PS   Claim 13; Fig 5; 89pp; English.
XX
CC   The present sequence is that of the heavy chain variable region (VH) of
CC   human single chain antibody (scFv) HIb-1 (see AAY95198), which is
CC   directed against platelet glycoprotein Ib (GPIb). The HIb series of scFv
CC   was isolated from a human synthetic VH and VL scFv library by 3 rounds of
CC   phagemid selection against transfected CHO cells expressing the GPIb
CC   alpha component of the GPIb/IX/V complex on their surface, followed by a
CC   4th round of selection against washed human platelets, and 2 final rounds
CC   in which attempts were made to displace scFv from washed platelets by
CC   flooding with murine monoclonal antibody or mimotope peptide (see
CC   AAY95229). Whether displayed as surface proteins on a phagemid or
CC   secreted as free scFv by Escherichia coli, the HIb scFv clones are
CC   capable of inhibiting von Willebrand factor-dependent aggregation of
CC   platelets. The scFv are composed of native human protein sequences and
CC   are therefore attractive potential reagents for therapeutic purposes.
CC   They provide a new class of antithrombotic agents, useful for the
CC   prevention of platelet-dependent thrombi in diseased arteries, bypass
CC   grafts, dialysis etc., and can also be used as diagnostic reagents.
CC   Methods of inhibiting aggregation of platelets, of binding human platelet
CC   GPIb alpha and of selecting a VH or VL region of an antibody that
CC   inhibits platelet aggregation are claimed
XX
SQ   Sequence 115 AA;

  Query Match             100.0%;  Score 96;  DB 3;  Length 115;
  Best Local Similarity   100.0%;  Pred. No. 8e-07;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         50 GINWNGGSTGYADSVKG 66


RESULT 11
AAY95190
ID   AAY95190 standard; protein; 115 AA.
XX
AC   AAY95190;
XX
DT   29-AUG-2000  (first entry)
XX
DE   Anti-platelet glycoprotein Ib human HIb-1 VH.
XX
KW   Variable heavy chain; single chain antibody; scFv; human; HIb-1;
KW   glycoprotein Ib alpha; platelet; aggregation; antiaggregant;
KW   antithrombotic; thrombus; therapy; diagnostic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .30
FT                   /note= "framework region 1"
FT   Region          31. .34
FT                   /note= "complementarity determining region 1"
FT   Region          35. .49
FT                   /note= "framework region 2"
FT   Region          50. .66
FT                   /note= "complementarity determining region 2"
FT   Region          67. .98
FT                   /note= "framework region 3"
FT   Region          99. .104
FT                   /note= "complementarity determining region 3"
FT   Region          105. .115
FT                   /note= "framework region 4"
XX
PN   WO200026667-A1.
XX
PD   11-MAY-2000.
XX
```

PF 29-OCT-1999; 99WO-US025495.
XX
PR 30-OCT-1998; 98US-0106275P.
XX
PA (MILL/) MILLER J L.
XX
PI Miller JL;
XX
DR WPI; 2000-365744/31.
XX
PT Isolated nucleic acid molecule encoding anti-human platelet glycoprotein
PT Ib alpha molecule useful for producing antibodies which inhibit platelet
PT aggregation.
XX
PS Claim 13; Page; 89pp; English.
XX
CC The present sequence is that of the heavy chain variable region (VH) of
CC human single chain antibody (scFv) HIb-1 (see AAY95198), which is
CC directed against platelet glycoprotein Ib (GPIb). The HIb series of scFv
CC was isolated from a human synthetic VH and VL scFv library by 3 rounds of
CC phagemid selection against transfected CHO cells expressing the GPIb
CC alpha component of the GPIb/IX/V complex on their surface, followed by a
CC 4th round of selection against washed human platelets, and 2 final rounds
CC in which attempts were made to displace scFv from washed platelets by
CC flooding with murine monoclonal antibody or mimotope peptide (see
CC AAY95229). Whether displayed as surface proteins on a phagemid or
CC secreted as free scFv by Escherichia coli, the HIb scFv clones are
CC capable of inhibiting von Willebrand factor-dependent aggregation of
CC platelets. The scFv are composed of native human protein sequences and
CC are therefore attractive potential reagents for therapeutic purposes.
CC They provide a new class of antithrombotic agents, useful for the
CC prevention of platelet-dependent thrombi in diseased arteries, bypass
CC grafts, dialysis etc., and can also be used as diagnostic reagents.
CC Methods of inhibiting aggregation of platelets, of binding human platelet
CC GPIb alpha and of selecting a VH or VL region of an antibody that
CC inhibits platelet aggregation are claimed. Note: The present sequence is
CC not shown in the specification but is derived from the HIb-1 VH sequence
CC given in Fig 5 (see AAY95189)
XX
SQ Sequence 115 AA;

Query Match 100.0%; Score 96; DB 3; Length 115;
Best Local Similarity 100.0%; Pred. No. 8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 50 GINWNGGSTGYADSVKG 66

RESULT 12
AAW19880
ID AAW19880 standard; protein; 116 AA.
XX
AC AAW19880;
XX
DT 07-DEC-1997 (first entry)
XX
DE CEA-specific antibody CEA5 VH sequence.
XX
KW Carcinoembryonic antigen; CEA; human; antibody; scFv; tumour marker;
KW lung cancer; breast cancer; colon cancer; adenocarcinoma; diagnosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 31. .35
FT /label= CDR1
FT /note= "complementarity determining region 1"
FT Region 50. .66
FT /label= CDR2
FT /note= "complementarity determining region 2"
FT Region 99. .105
FT /label= CDR3
FT /note= "complementarity determining region 3"
XX
PN WO9720932-A1.
XX
PD 12-JUN-1997.
XX
PF 09-DEC-1996; 96WO-GB003043.
XX
PR 07-DEC-1995; 95GB-00025004.
PR 23-MAY-1996; 96GB-00010824.
PR 11-OCT-1996; 96GB-00021295.
XX
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
PI Osbourn JK, Allen DJ, Mccafferty JG;
XX
DR WPI; 1997-319779/29.
DR N-PSDB; AAT72130.
XX
PT Specific binding members for human carcinoembryonic antigen - bind to the
PT A3-B3 extracellular domain of hCEA and are substantially non-cross-
PT reactive with human liver cells; used for diagnosing cancer.
XX
PS Claim 7; Fig 1a; 128pp; English.
XX
CC This polypeptide sequence comprises the heavy chain variable region (VH)
CC of human carcinoembryonic antigen (hCEA)-specific antibody CEA5. VH
CC (AAT72126-32) and VL (AAT72133-35) gene sequences were obtained for anti-
CC hCEA antibodies CEA1-CEA7 (see AAW19876-85). A claimed specific binding
CC member (A) comprises an hCEA specific antibody antigen binding domain
CC that has a dissociation constant for hCEA of less than 1 x 10 -8 M, is
CC non-cross-reactive with human liver cells, and preferentially binds to
CC the A3-B3 extracellular domain of hCEA and/or to cell-associated hCEA
CC over hCEA over soluble hCEA. Preferred (A) include pairings of VH and VL
CC sequences from CEA1-7, or their CDR sequences, as well as CEA6 VH and VL
CC variants. (A) is used to detect cells expressing hCEA, in vivo or in
CC vitro, especially tumour cells for diagnosing cancer, e.g. adenocarcinoma
CC of the colon, lung or breast
XX
SQ Sequence 116 AA;

Query Match 100.0%; Score 96; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 50 GINWNGGSTGYADSVKG 66

RESULT 13
AAR66312
ID AAR66312 standard; protein; 117 AA.
XX
AC AAR66312;
XX
DT 25-MAR-2003 (revised)
DT 03-AUG-1995 (first entry)
XX
DE Human immunoglobulin variable heavy chain #18.
XX
KW Primer; PCR; amplify; human; immunoglobulin; variable; heavy chain;
KW cosmid; placenta; vector; pJB81; E.coli; mammalian.
XX
OS Homo sapiens.
XX
PN WO9426895-A1.
XX
PD 24-NOV-1994.
XX
PF 10-MAY-1993; 93WO-JP000603.
XX

```
PR   10-MAY-1993;   93WO-JP000603.
XX
PA   (NISB ) JAPAN TOBACCO INC.
XX
PI   Honjo T,  Matsuda F;
XX
DR   WPI; 1995-006791/01.
DR   N-PSDB; AAQ78958.
XX
PT   DNA fragment comprising human immunoglobulin Vh genes - for the
PT   production of human immunoglobulin in mammalian hosts.
XX
PS   Claim 29; Page 56-57; 130pp; Japanese.
XX
CC   Protein sequences (AAR66295-51) are novel human immunoglobulin heavy
CC   chain sequences encoded by novel isolated genes. The genes (AAQ78939-
CC   79002) were isolated and cloned from a series of cosmid constructs: Y202;
CC   Y103; Y21; Y6;Y24; 3-31; M84; M118 and M131, by PCR amplification using
CC   primers AAQ78917-38. The genes are subdivided into 5 families of Vh
CC   genes. The fragments cover a region of 800 kb. The DNA fragments were
CC   isolated from high molecular weight DNA from human placenta. The DNA was
CC   partially digested with TaqI restriction enzyme. The fragments were
CC   separated by gel elecrophoresis and 35-45 kb fractions were collected.
CC   The fragments were ligated with ClaI-digested cosmid vector pJB81. The
CC   ligation products were in vitro packed and infected into E.coli 490A. The
CC   fragments were then subcloned by colony hybridisation. The Vh genes and
CC   the DNA fragments encoding them are useful in producing human
CC   immunoglobulin in mammalian hosts. (Updated on 25-MAR-2003 to correct PN
CC   field.)
XX
SQ   Sequence 117 AA;

  Query Match             100.0%;  Score 96;  DB 2;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-07;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         69 GINWNGGSTGYADSVKG 85


RESULT 14
AAU02560
ID   AAU02560 standard; protein; 118 AA.
XX
AC   AAU02560;
XX
DT   29-AUG-2001  (first entry)
XX
DE   Anti-adipocyte monoclonal antibody heavy chain, FAT 46.
XX
KW   Antibody; adipocyte; heavy chain; light chain; obesity; fat;
KW   heart disease; complementarity determining region; CDR.
XX
OS   Homo sapiens.
XX
PN   WO200127279-A1.
XX
PD   19-APR-2001.
XX
PF   11-OCT-2000; 2000WO-GB003900.
XX
PR   12-OCT-1999;   99US-0158812P.
XX
PA   (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
PI   Edwards BM,  Main SH,  Vaughan TJ;
XX
DR   WPI; 2001-282031/29.
DR   N-PSDB; AAS03460.
XX
PT   Panel of specific binding members of antibody molecules which bind to
PT   whole adipocytes is used in the treatment of obesity and obesity related
PT   diseases.
XX
PS   Claim 1; Page 130; 182pp; English.
XX
CC   AAU02501-AAU02635, and AAU02641-AAU02748 represent the amino acid
CC   sequences of anti-adipocyte monoclonal antibody heavy chain, light chain,
CC   and heavy chain complementarity determining regions (CDR) of the
CC   invention. The antibodies can be used in the treatment of obesity and
CC   obesity related diseases. The antibodies can be used to deliver drugs or
CC   pro-drugs directly to the fat mass of an obese patient or the antibody
CC   can be used as a therapeutic itself. Antibodies binding specifically to
CC   adipocytes can be used to activate the immune system to destroy the cells
CC   by complement mediated lysis. The antibodies may be labeled with a
CC   detectable label such as radiolabel, fluorescent or chemical group and
CC   used in methods of diagnosis in human subjects e.g. to determine the
CC   presence of adipocyte antigen on the surface of an adipocyte to detect or
CC   determine the presence or level of adipocytes in a cell or tissue sample.
CC   The antibodies can be used as an alternative means of treatment for obese
CC   patients other than undergoing surgery to remove excess fat. Antibodies
CC   for different types of fat deposits can also be produced e.g. intra-
CC   abdominal fat associated with heart disease
XX
SQ   Sequence 118 AA;

  Query Match             100.0%;  Score 96;  DB 4;  Length 118;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-07;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         50 GINWNGGSTGYADSVKG 66


RESULT 15
AAY96065
ID   AAY96065 standard; protein; 122 AA.
XX
AC   AAY96065;
XX
DT   05-DEC-2000  (first entry)
XX
DE   Human anti-DAF antibody LU20 heavy chain variable region.
XX
KW   LU20; human; antibody; VH domain; decay accelerating factor; DAF;
KW   phage display; subtractive panning; lung cancer; lung carcinoma;
KW   lung adenocarcinoma; therapy; diagnosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          26. .32
FT                   /note= "hypervariable loop region"
FT   Region          31. .35
FT                   /label= CDRI
FT                   /note= "complementarity determining region I"
FT   Region          50. .66
FT                   /label= CDRII
FT                   /note= "complementarity determining region II"
FT   Region          53. .56
FT                   /note= "hypervariable loop residues"
FT   Region          99. .111
FT                   /label= CDRIII
FT                   /note= "complementarity determining region III"
FT   Region          100. .110
FT                   /note= "hypervariable loop residues"
XX
PN   WO200052054-A2.
XX
PD   08-SEP-2000.
XX
PF   29-FEB-2000; 2000WO-US005352.
XX
PR   01-MAR-1999;   99US-0122262P.
```

XX
PA (GETH) GENENTECH INC.
XX
PI Carter PJ, Ridgway JB;
XX
DR WPI; 2000-594169/56.
XX
PT Making antibodies (e.g. anti-decay accelerating factor antibody) for
PT diagnosing or treating e.g. lung cancer comprises identifying an antigen
PT that is differentially expressed on the surface of two or more distinct
PT cell populations.
XX
PS Disclosure; Fig 5B; 52pp; English.
XX
CC The present sequence is that of the heavy chain variable region (VH) of
CC the anti-decay accelerating factor (DAF) human antibody LU20. The VL
CC region is given in AAY06062. LU20 was produced using a novel method for
CC making antibodies which can be used for cancer diagnosis or therapy. The
CC method comprises: (a) binding an antibody phage from a naive antibody
CC phage library to a live cancer cell; (b) selecting an antibody phage or
CC antibody which binds selectively to the live cancer cell; and (c)
CC identifying an antigen to which the antibody phage or antibody binds. To
CC obtain LU20, a human scFv library was used to search for tumour-
CC associated antigens by panning the lung adenocarcinoma cell line 1264,
CC and counter-selecting with a non-tumour bronchial epithelial cell line,
CC BEAS-2B. The invention also describes a method for identifying an antigen
CC which is differentially expressed on the surface of 2 or more distinct
CC cell populations. The anti-DAF human antibody, or a composition
CC comprising the antibody, is useful for in vivo cancer diagnosis or
CC therapy. In particular, the antibody is useful for diagnosing or treating
CC lung cancer, e.g. small-cell lung cancer, non-small cell lung cancer,
CC large cell lung carcinoma, lung adenocarcinoma, or squamous cell lung
CC carcinoma (all claimed)
XX
SQ Sequence 122 AA;

Query Match 100.0%; Score 96; DB 3; Length 122;
Best Local Similarity 100.0%; Pred. No. 8.5e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         50 GINWNGGSTGYADSVKG 66
```

Search completed: September 18, 2004, 03:42:54
Job time : 121.214 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 03:39:30 ; Search time 39.4643 Seconds
(without alignments)
22.239 Million cell updates/sec

Title: US-10-029-926B-115
Perfect score: 96
Sequence: 1 GINWNGGSTGYADSVKG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	96	100.0	116	2	US-08-652-816A-14	Sequence 14, Appl
2	96	100.0	117	3	US-08-545-809A-106	Sequence 106, App
3	96	100.0	309	4	US-09-079-029-9	Sequence 9, Appli
4	79	82.3	118	3	US-08-545-809A-97	Sequence 97, Appl
5	79	82.3	120	2	US-08-958-201-8	Sequence 8, Appli
6	79	82.3	120	2	US-08-958-201-10	Sequence 10, Appl
7	71	74.0	117	3	US-08-545-809A-119	Sequence 119, App
8	71	74.0	118	3	US-08-545-809A-125	Sequence 125, App
9	68	70.8	117	1	US-07-942-245-24	Sequence 24, Appl
10	63	65.6	126	3	US-08-983-607-26	Sequence 26, Appl
11	62.5	65.1	117	3	US-09-157-370-1	Sequence 1, Appli
12	62	64.6	17	1	US-08-264-093-22	Sequence 22, Appl
13	62	64.6	111	4	US-09-899-896-7	Sequence 7, Appli
14	62	64.6	118	2	US-08-652-816A-11	Sequence 11, Appl
15	62	64.6	120	1	US-08-264-093-14	Sequence 14, Appl
16	62	64.6	483	3	US-09-049-672A-5	Sequence 5, Appli
17	61.5	64.1	95	3	US-09-043-514-2	Sequence 2, Appli
18	61.5	64.1	116	3	US-08-545-809A-135	Sequence 135, App
19	61	63.5	113	3	US-08-974-899-6	Sequence 6, Appli
20	61	63.5	116	3	US-09-184-658-48	Sequence 48, Appl
21	61	63.5	116	3	US-08-983-607-36	Sequence 36, Appl
22	61	63.5	116	4	US-09-504-262D-48	Sequence 48, Appl
23	61	63.5	117	3	US-08-545-809A-103	Sequence 103, App
24	61	63.5	117	3	US-08-545-809A-109	Sequence 109, App
25	61	63.5	117	3	US-08-983-607-46	Sequence 46, Appl
26	61	63.5	118	2	US-08-652-816A-12	Sequence 12, Appl
27	61	63.5	119	4	US-09-648-067A-15	Sequence 15, Appl
28	61	63.5	120	2	US-08-428-197-20	Sequence 20, Appl
29	61	63.5	120	2	US-08-428-197-22	Sequence 22, Appl
30	61	63.5	120	2	US-08-428-197-24	Sequence 24, Appl
31	61	63.5	120	2	US-08-428-197-26	Sequence 26, Appl
32	61	63.5	120	2	US-08-428-197-28	Sequence 28, Appl
33	61	63.5	120	2	US-08-428-197-30	Sequence 30, Appl
34	61	63.5	120	2	US-08-428-197-32	Sequence 32, Appl
35	61	63.5	120	2	US-08-428-197-34	Sequence 34, Appl
36	61	63.5	120	2	US-08-428-197-40	Sequence 40, Appl
37	61	63.5	120	4	US-09-025-769B-38	Sequence 38, Appl
38	61	63.5	120	4	US-09-025-769B-63	Sequence 63, Appl
39	61	63.5	120	5	PCT-US93-07832-4	Sequence 4, Appli
40	61	63.5	120	5	PCT-US93-10555-20	Sequence 20, Appl
41	61	63.5	120	5	PCT-US93-10555-22	Sequence 22, Appl
42	61	63.5	120	5	PCT-US93-10555-24	Sequence 24, Appl
43	61	63.5	120	5	PCT-US93-10555-26	Sequence 26, Appl
44	61	63.5	120	5	PCT-US93-10555-28	Sequence 28, Appl
45	61	63.5	120	5	PCT-US93-10555-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-08-652-816A-14
; Sequence 14, Application US/08652816A
; Patent No. 5872215
; GENERAL INFORMATION:
; APPLICANT: Osbourn, JK
; APPLICANT: Allen, DJ
; APPLICANT: McCafferty, JG
; TITLE OF INVENTION: Specific binding members, materials and
; TITLE OF INVENTION: methods.
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,816A
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.4
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.8
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206318.9
; FILING DATE: 24-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206372.6
; FILING DATE: 23-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9525004.9
; FILING DATE: 07-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610824.6
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/02240
; FILING DATE: 02-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/244,597
; FILING DATE: 01-JUN-1994
; ATTORNEY/AGENT INFORMATION:

```
;      NAME:  David W. Clough
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  28111/33308
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  116 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-652-816A-14

  Query Match             100.0%;  Score 96;  DB 2;  Length 116;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-06;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         50 GINWNGGSTGYADSVKG 66


RESULT 2
US-08-545-809A-106
; Sequence 106, Application US/08545809A
; Patent No. 6096878
;  GENERAL INFORMATION:
;    APPLICANT:  Honjo, Tasuku
;    APPLICANT:  Matsuda, Fumihiko
;    TITLE OF INVENTION:  HUMAN IMMUNOGLOBULIN VH GENE
;    TITLE OF INVENTION:  SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
;    NUMBER OF SEQUENCES:  145
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/545,809A
;      FILING DATE:  27-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP93/00603
;      FILING DATE:  10-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Freeman, John W.
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  06501/004001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-5070
;      TELEFAX:  617-542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  117 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-545-809A-106

  Query Match             100.0%;  Score 96;  DB 3;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-06;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         69 GINWNGGSTGYADSVKG 85


RESULT 3
US-09-079-029-9
; Sequence 9, Application US/09079029
; Patent No. 6342369
;  GENERAL INFORMATION:
;    APPLICANT:  Adams, Camilia W.
;    APPLICANT:  Ashkenazi, Avi J.
;    APPLICANT:  Chuntharapai, Anan
;    APPLICANT:  Kim, Kyung J.
;    TITLE OF INVENTION:  Apo-2 Receptor
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  1 DNA Way
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,029
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Marschang, Diane L.
;      REGISTRATION NUMBER:  35,600
;      REFERENCE/DOCKET NUMBER:  P1101R2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650/225-5416
;      TELEFAX:  650/952-9881
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  309 amino acids
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
US-09-079-029-9

  Query Match             100.0%;  Score 96;  DB 4;  Length 309;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-06;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         89 GINWNGGSTGYADSVKG 105


RESULT 4
US-08-545-809A-97
; Sequence 97, Application US/08545809A
; Patent No. 6096878
;  GENERAL INFORMATION:
;    APPLICANT:  Honjo, Tasuku
;    APPLICANT:  Matsuda, Fumihiko
;    TITLE OF INVENTION:  HUMAN IMMUNOGLOBULIN VH GENE
;    TITLE OF INVENTION:  SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
;    NUMBER OF SEQUENCES:  145
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/545,809A
;      FILING DATE:  27-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP93/00603
;      FILING DATE:  10-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Freeman, John W.
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  06501/004001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-5070
;      TELEFAX:  617-542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  97:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  118 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-545-809A-97

  Query Match             82.3%;  Score 79;  DB 3;  Length 118;
  Best Local Similarity   82.4%;  Pred. No. 0.00043;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||:|| || ||||||||
Db         69 GISWNSGSIGYADSVKG 85


RESULT 5
US-08-958-201-8
; Sequence 8, Application US/08958201
; Patent No. 5977319
;  GENERAL INFORMATION:
;    APPLICANT:  Pope, Anthony R
;    APPLICANT:  Pritchard, Kevin
;    APPLICANT:  Williams, Andrew J
;    APPLICANT:  Johnson, Kevin S
;    TITLE OF INVENTION:  Specific binding members for estradiol;
;    TITLE OF INVENTION:  materials and methods
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall O'Toole Gerstein Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/958,201
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,897
;      FILING DATE:  21-OCT-1996
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  120 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  2D
US-08-958-201-8

  Query Match             82.3%;  Score 79;  DB 2;  Length 120;
  Best Local Similarity   82.4%;  Pred. No. 0.00044;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||:|| || ||||||||
Db         50 GISWNSGSIGYADSVKG 66


RESULT 6
US-08-958-201-10
; Sequence 10, Application US/08958201
; Patent No. 5977319
;  GENERAL INFORMATION:
;    APPLICANT:  Pope, Anthony R
;    APPLICANT:  Pritchard, Kevin
;    APPLICANT:  Williams, Andrew J
;    APPLICANT:  Johnson, Kevin S
;    TITLE OF INVENTION:  Specific binding members for estradiol;
;    TITLE OF INVENTION:  materials and methods
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall O'Toole Gerstein Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/958,201
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,897
;      FILING DATE:  21-OCT-1996
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  120 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  2Db
US-08-958-201-10

  Query Match             82.3%;  Score 79;  DB 2;  Length 120;
  Best Local Similarity   82.4%;  Pred. No. 0.00044;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||:|| || ||||||||
Db         50 GISWNSGSIGYADSVKG 66


RESULT 7
US-08-545-809A-119
; Sequence 119, Application US/08545809A
; Patent No. 6096878
;  GENERAL INFORMATION:
;    APPLICANT:  Honjo, Tasuku
;    APPLICANT:  Matsuda, Fumihiko
;    TITLE OF INVENTION:  HUMAN IMMUNOGLOBULIN VH GENE
;    TITLE OF INVENTION:  SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
;    NUMBER OF SEQUENCES:  145
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
```

```
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  Windows95
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/545,809A
;       FILING DATE:  27-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/JP93/00603
;       FILING DATE:  10-MAY-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Freeman, John W.
;       REGISTRATION NUMBER:  29,066
;       REFERENCE/DOCKET NUMBER:  06501/004001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-542-5070
;       TELEFAX:  617-542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  119:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  117 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-545-809A-119

  Query Match             74.0%;  Score 71;  DB 3;  Length 117;
  Best Local Similarity   70.6%;  Pred. No. 0.0056;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |::|||   | |||||||
Db         69 GVSWNGSRTHYADSVKG 85


RESULT 8
US-08-545-809A-125
; Sequence 125, Application US/08545809A
; Patent No. 6096878
;  GENERAL INFORMATION:
;    APPLICANT:  Honjo, Tasuku
;    APPLICANT:  Matsuda, Fumihiko
;    TITLE OF INVENTION:  HUMAN IMMUNOGLOBULIN VH GENE
;    TITLE OF INVENTION:  SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
;    NUMBER OF SEQUENCES:  145
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/545,809A
;      FILING DATE:  27-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP93/00603
;      FILING DATE:  10-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Freeman, John W.
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  06501/004001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-5070
;      TELEFAX:  617-542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  125:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  118 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-545-809A-125

  Query Match             74.0%;  Score 71;  DB 3;  Length 118;
  Best Local Similarity   81.2%;  Pred. No. 0.0057;
  Matches   13;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              |:|:|||| |||||||
Db         70 ISWDGGSTYYADSVKG 85


RESULT 9
US-07-942-245-24
; Sequence 24, Application US/07942245
; Patent No. 5639641
;  GENERAL INFORMATION:
;    APPLICANT:  PEDERSEN, Jan T.
;    APPLICANT:  SEARLE,   Stephen M.J.
;    APPLICANT:  REES,     Anthony R.
;    APPLICANT:  ROGUSKA,  Michael A.
;    APPLICANT:  GUILD,    Braydon C.
;    TITLE OF INVENTION:  SURFACE RESIDUE VENEERING OF RODENT
;    TITLE OF INVENTION:  ANTIBODIES
;    NUMBER OF SEQUENCES:  522
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak & Seas
;      STREET:  2100 Pensylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  HP 9000/700 Workstation
;      OPERATING SYSTEM:  UNIX
;      SOFTWARE:  In house
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/942,245
;      FILING DATE:  09-SEP-1992
;      CLASSIFICATION:  530
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  117 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-942-245-24

  Query Match             70.8%;  Score 68;  DB 1;  Length 117;
  Best Local Similarity   70.6%;  Pred. No. 0.015;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||:|:  | ||||||||
Db         50 GISWDSSSIGYADSVKG 66


RESULT 10
US-08-983-607-26
; Sequence 26, Application US/08983607
; Patent No. 6140470
;  GENERAL INFORMATION:
;    APPLICANT:  Alan Garen
;    APPLICANT:  Xiaohong Cai
;    TITLE OF INVENTION:  Human Anti-Tumor Monoclonal Anti-
;    TITLE OF INVENTION:  bodies
```

```
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Department of Molecular Biophysics
;      ADDRESSEE:  and Biochemistry, Yale University
;      STREET:  266 Whitney Avenue
;      CITY:  New Haven
;      STATE:  Connecticut
;      COUNTRY:  United States of America
;      ZIP:  06520-8114
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" 1.44 Mb diskette
;      COMPUTER:  IBM PC
;      OPERATING SYSTEM:  MS DOS
;      SOFTWARE:  Word Processing
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/983,607
;      FILING DATE:  April 27, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/IB96/01032
;      FILING DATE:  June 28, 1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary M. Krinsky
;      REGISTRATION NUMBER:  32423
;      REFERENCE/DOCKET NUMBER:  OCR-679
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  203-773-9544
;      TELEFAX:  203-773-1183
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  126 residues
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  polypeptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens (melanoma patient immu-
;      ORGANISM:  nized with autologous tumor cells)
;      INDIVIDUAL ISOLATE:  peripheral blood lymphocytes
;    IMMEDIATE SOURCE:
;      LIBRARY:  DM414 scFv antibodies obtained from
;      LIBRARY:  fUSE5 fusion phage construct
;      CLONE:  Z78
;    FEATURE:
;      NAME/KEY:  heavy chain
US-08-983-607-26

  Query Match             65.6%;  Score 63;  DB 3;  Length 126;
  Best Local Similarity   81.2%;  Pred. No. 0.081;
  Matches   13;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              |: ||||| |||||||
Db         51 ISGNGGSTYYADSVKG 66


RESULT 11
US-09-157-370-1
; Sequence 1, Application US/09157370A
; Patent No. 6262238
; GENERAL INFORMATION:
;  APPLICANT: STEIPE, Boris
;  APPLICANT: STEINBACHER, Stefan
;  TITLE OF INVENTION: PROCESS FOR MODIFYING THE STABILITY OF ANTIBODIES
;  FILE REFERENCE: P8341-8072
;  CURRENT APPLICATION NUMBER: US/09/157,370A
;  CURRENT FILING DATE: 1998-09-21
;  EARLIER APPLICATION NUMBER: 08/765,179
;  EARLIER FILING DATE: 1997-01-14
;  EARLIER APPLICATION NUMBER: PCT/EP95/02626
;  EARLIER FILING DATE: 1995-07-06
;  EARLIER APPLICATION NUMBER: DE/P44 25 115.7
;  EARLIER FILING DATE: 1994-07-15
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 117
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-157-370-1

  Query Match             65.1%;  Score 62.5;  DB 3;  Length 117;
  Best Local Similarity   59.1%;  Pred. No. 0.088;
  Matches   13;  Conservative    1;  Mismatches    3;  Indels    5;  Gaps    1;

Qy          1 GINW-----NGGSTGYADSVKG 17
              |: |     ||| | |||||||
Db         44 GLEWVGWIYNGGDTYYADSVKG 65


RESULT 12
US-08-264-093-22
; Sequence 22, Application US/08264093
; Patent No. 5639863
;  GENERAL INFORMATION:
;    APPLICANT:  Michael D. Dan
;    TITLE OF INVENTION:  HUMAN MONOCLONAL ANTIBODIES SPECIFIC TO
;    TITLE OF INVENTION:  CELL CYCLE-INDEPENDENT GLIOMA SURFACE
;    TITLE OF INVENTION:  ANTIGEN
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ridout & Maybee
;      STREET:  2300 Richmond-Adelaide Centre
;      STREET:  101 Richmond Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 2J7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.4 Mb storage
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  MS-DOS 6.00
;      SOFTWARE:  ASCII Editor
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/264,093
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:  No. 5639863 applicable
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lake, James R.
;      REGISTRATION NUMBER:  31081
;      REFERENCE/DOCKET NUMBER:  NOVOP/106A/7551
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 868-1482
;      TELEFAX:  (416) 362-0823
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not applicable
;      TOPOLOGY:  linear
US-08-264-093-22

  Query Match             64.6%;  Score 62;  DB 1;  Length 17;
  Best Local Similarity   81.2%;  Pred. No. 0.014;
  Matches   13;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              |: ||||| |||||||
Db          2 ISSNGGSTYYADSVKG 17


RESULT 13
US-09-899-896-7
```

```
; Sequence 7, Application US/09899896
; Patent No. 6569431
; GENERAL INFORMATION:
;  APPLICANT: von Bdingen, Hans-Christian
;  APPLICANT:  Genain, Claude P.
;  APPLICANT:  Hauser, Stephen L.
;  TITLE OF INVENTION: Recombinant Antibody Fragments as Autoantibody
;  TITLE OF INVENTION:  Antagonists
;  FILE REFERENCE: SF01-025-2
;  CURRENT APPLICATION NUMBER: US/09/899,896
;  CURRENT FILING DATE:  2001-07-05
;  PRIOR APPLICATION NUMBER: 09/691,654
;  PRIOR FILING DATE: 2000-10-17
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 111
;   TYPE: PRT
;   ORGANISM: human
US-09-899-896-7

  Query Match             64.6%;  Score 62;  DB 4;  Length 111;
  Best Local Similarity   75.0%;  Pred. No. 0.098;
  Matches   12;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              |:::|||| |||||||
Db         51 ISYDGGSTYYADSVKG 66


RESULT 14
US-08-652-816A-11
; Sequence 11, Application US/08652816A
; Patent No. 5872215
;  GENERAL INFORMATION:
;    APPLICANT:  Osbourn, JK
;    APPLICANT:  Allen, DJ
;    APPLICANT:  McCafferty, JG
;    TITLE OF INVENTION:  Specific binding members, materials and
;    TITLE OF INVENTION:  methods.
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,816A
;      FILING DATE:  23-MAY-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9125579.4
;      FILING DATE:  02-DEC-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9125579.8
;      FILING DATE:  02-DEC-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9206318.9
;      FILING DATE:  24-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9206372.6
;      FILING DATE:  23-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9525004.9
;      FILING DATE:  07-DEC-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9610824.6
;      FILING DATE:  23-MAY-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/02240
;      FILING DATE:  02-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/244,597
;      FILING DATE:  01-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  David W. Clough
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  28111/33308
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  118 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-652-816A-11

  Query Match             64.6%;  Score 62;  DB 2;  Length 118;
  Best Local Similarity   81.2%;  Pred. No. 0.1;
  Matches   13;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              |: ||||| |||||||
Db         51 ISSNGGSTYYADSVKG 66


RESULT 15
US-08-264-093-14
; Sequence 14, Application US/08264093
; Patent No. 5639863
;  GENERAL INFORMATION:
;    APPLICANT:  Michael D. Dan
;    TITLE OF INVENTION:  HUMAN MONOCLONAL ANTIBODIES SPECIFIC TO
;    TITLE OF INVENTION:  CELL CYCLE-INDEPENDENT GLIOMA SURFACE
;    TITLE OF INVENTION:  ANTIGEN
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ridout & Maybee
;      STREET:  2300 Richmond-Adelaide Centre
;      STREET:  101 Richmond Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 2J7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.4 Mb storage
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  MS-DOS 6.00
;      SOFTWARE:  ASCII Editor
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/264,093
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:  No. 5639863 applicable
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lake, James R.
;      REGISTRATION NUMBER:  31081
;      REFERENCE/DOCKET NUMBER:  NOVOP/106A/7551
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 868-1482
;      TELEFAX:  (416) 362-0823
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  120 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not applicable
;      TOPOLOGY:  linear
US-08-264-093-14

  Query Match             64.6%;  Score 62;  DB 1;  Length 120;
  Best Local Similarity   81.2%;  Pred. No. 0.11;
```

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          2 INWNGGSTGYADSVKG 17
              |: ||||| |||||||
Db         44 ISSNGGSTYYADSVKG 59
```

Search completed: September 18, 2004, 03:48:47
Job time : 40.4643 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 03:46:39 ; Search time 675.143 Seconds
(without alignments)
8.086 Million cell updates/sec

Title: US-10-029-926B-115
Perfect score: 96
Sequence: 1 GINWNGGSTGYADSVKG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	96	100.0	17	12	US-10-029-926B-115	Sequence 115, App
2	96	100.0	17	15	US-10-032-037B-115	Sequence 115, App
3	96	100.0	17	15	US-10-029-988B-115	Sequence 115, App
4	96	100.0	17	15	US-10-032-423A-115	Sequence 115, App
5	96	100.0	98	12	US-10-453-698-60	Sequence 60, Appl
6	96	100.0	98	12	US-10-029-926B-61	Sequence 61, Appl
7	96	100.0	98	14	US-10-194-975-20	Sequence 20, Appl
8	96	100.0	98	15	US-10-308-817-60	Sequence 60, Appl
9	96	100.0	98	15	US-10-032-037B-61	Sequence 61, Appl
10	96	100.0	98	15	US-10-029-988B-61	Sequence 61, Appl
11	96	100.0	98	15	US-10-032-423A-61	Sequence 61, Appl
12	96	100.0	98	16	US-10-379-392-21	Sequence 21, Appl
13	96	100.0	122	15	US-10-447-331-6	Sequence 6, Appli
14	96	100.0	239	10	US-09-880-748-937	Sequence 937, App
15	96	100.0	239	10	US-09-880-748-2015	Sequence 2015, Ap
16	96	100.0	239	10	US-09-880-748-2038	Sequence 2038, Ap
17	96	100.0	239	12	US-10-293-418-937	Sequence 937, App
18	96	100.0	239	12	US-10-293-418-2015	Sequence 2015, Ap
19	96	100.0	239	12	US-10-293-418-2038	Sequence 2038, Ap
20	96	100.0	244	14	US-10-322-673-42	Sequence 42, Appl
21	96	100.0	266	15	US-10-032-037B-204	Sequence 204, App
22	96	100.0	266	15	US-10-029-988B-204	Sequence 204, App
23	96	100.0	266	15	US-10-032-423A-204	Sequence 204, App
24	96	100.0	277	12	US-10-029-926B-25	Sequence 25, Appl
25	96	100.0	277	12	US-10-029-926B-203	Sequence 203, App
26	96	100.0	277	15	US-10-032-037B-25	Sequence 25, Appl
27	96	100.0	277	15	US-10-032-037B-203	Sequence 203, App
28	96	100.0	277	15	US-10-029-988B-25	Sequence 25, Appl
29	96	100.0	277	15	US-10-029-988B-203	Sequence 203, App
30	96	100.0	277	15	US-10-032-423A-25	Sequence 25, Appl
31	96	100.0	277	15	US-10-032-423A-203	Sequence 203, App
32	96	100.0	309	13	US-10-052-798-9	Sequence 9, Appli
33	96	100.0	309	14	US-10-288-917-9	Sequence 9, Appli
34	96	100.0	309	15	US-10-423-448-9	Sequence 9, Appli
35	96	100.0	464	12	US-10-029-926B-26	Sequence 26, Appl
36	96	100.0	464	15	US-10-032-037B-26	Sequence 26, Appl
37	96	100.0	464	15	US-10-029-988B-26	Sequence 26, Appl
38	96	100.0	464	15	US-10-032-423A-26	Sequence 26, Appl
39	79	82.3	98	12	US-10-029-926B-60	Sequence 60, Appl
40	79	82.3	98	15	US-10-032-037B-60	Sequence 60, Appl
41	79	82.3	98	15	US-10-029-988B-60	Sequence 60, Appl
42	79	82.3	98	15	US-10-032-423A-60	Sequence 60, Appl
43	79	82.3	99	12	US-10-453-698-56	Sequence 56, Appl
44	79	82.3	99	14	US-10-194-975-16	Sequence 16, Appl
45	79	82.3	99	15	US-10-308-817-56	Sequence 56, Appl

ALIGNMENTS

RESULT 1
US-10-029-926B-115
; Sequence 115, Application US/10029926B
; Publication No. US20040073011A1
; GENERAL INFORMATION:
; APPLICANT: HAGAY, et al.
; TITLE OF INVENTION: SPECIFIC HUMAN ANTIBODIES FOR SELECTIVE CANCER THERAPY
; FILE REFERENCE: 10793/50
; CURRENT APPLICATION NUMBER: US/10/029,926B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 115
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-926B-115

Query Match 100.0%; Score 96; DB 12; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
|||||||||||||||||
Db 1 GINWNGGSTGYADSVKG 17

RESULT 2
US-10-032-037B-115
; Sequence 115, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/44

; CURRENT APPLICATION NUMBER: US/10/032,037B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 115
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-037B-115

Query Match 100.0%; Score 96; DB 15; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 1 GINWNGGSTGYADSVKG 17

RESULT 3
US-10-029-988B-115
; Sequence 115, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/46
; CURRENT APPLICATION NUMBER: US/10/029,988B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 115
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-988B-115

Query Match 100.0%; Score 96; DB 15; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 1 GINWNGGSTGYADSVKG 17

RESULT 4
US-10-032-423A-115
; Sequence 115, Application US/10032423A
; Publication No. US20040002450A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/45
; CURRENT APPLICATION NUMBER: US/10/032,423A
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 115
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-423A-115

Query Match 100.0%; Score 96; DB 15; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 1 GINWNGGSTGYADSVKG 17

RESULT 5
US-10-453-698-60
; Sequence 60, Application US/10453698
; Publication No. US20040038308A1
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 82 CIP (1087-37 CIP)
; CURRENT APPLICATION NUMBER: US/10/453,698
; CURRENT FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 60
; LENGTH: 98
; TYPE: PRT
; ORGANISM: human
US-10-453-698-60

Query Match 100.0%; Score 96; DB 12; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 50 GINWNGGSTGYADSVKG 66

RESULT 6
US-10-029-926B-61
; Sequence 61, Application US/10029926B
; Publication No. US20040073011A1
; GENERAL INFORMATION:
; APPLICANT: HAGAY, et al.
; TITLE OF INVENTION: SPECIFIC HUMAN ANTIBODIES FOR SELECTIVE CANCER THERAPY
; FILE REFERENCE: 10793/50
; CURRENT APPLICATION NUMBER: US/10/029,926B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-926B-61

Query Match 100.0%; Score 96; DB 12; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 50 GINWNGGSTGYADSVKG 66

RESULT 7
US-10-194-975-20
; Sequence 20, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975

; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-194-975-20

Query Match 100.0%; Score 96; DB 14; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 50 GINWNGGSTGYADSVKG 66

RESULT 8
US-10-308-817-60
; Sequence 60, Application US/10308817
; Publication No. US20030219861A1
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; APPLICANT: Wu, Dayang
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 1087-37
; CURRENT APPLICATION NUMBER: US/10/308,817
; CURRENT FILING DATE: 2002-12-03
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 98
; TYPE: PRT
; ORGANISM: human
US-10-308-817-60

Query Match 100.0%; Score 96; DB 15; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 50 GINWNGGSTGYADSVKG 66

RESULT 9
US-10-032-037B-61
; Sequence 61, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/44
; CURRENT APPLICATION NUMBER: US/10/032,037B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-037B-61

Query Match 100.0%; Score 96; DB 15; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 50 GINWNGGSTGYADSVKG 66

RESULT 10
US-10-029-988B-61
; Sequence 61, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/46
; CURRENT APPLICATION NUMBER: US/10/029,988B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-988B-61

Query Match 100.0%; Score 96; DB 15; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 50 GINWNGGSTGYADSVKG 66

RESULT 11
US-10-032-423A-61
; Sequence 61, Application US/10032423A
; Publication No. US20040002450A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/45
; CURRENT APPLICATION NUMBER: US/10/032,423A
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-423A-61

Query Match 100.0%; Score 96; DB 15; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
||||||||||||||||||
Db 50 GINWNGGSTGYADSVKG 66

RESULT 12
US-10-379-392-21
; Sequence 21, Application US/10379392
; Publication No. US20040110226A1
; GENERAL INFORMATION:
; APPLICANT: Lazar, Gregory Alan
; APPLICANT: Desjarlais, John Rudolf
; APPLICANT: Marshall, Shannon Alicia

```
;  APPLICANT:  Dahiyat, Bassil I.
;  TITLE OF INVENTION: ANTIBODY OPTIMIZATION
;  FILE REFERENCE: A-71386-3 463077-236
;  CURRENT APPLICATION NUMBER: US/10/379,392
;  CURRENT FILING DATE:  2003-03-03
;  PRIOR APPLICATION NUMBER: US 60/360,843
;  PRIOR FILING DATE: 2002-03-01
;  PRIOR APPLICATION NUMBER: US 60/384,197
;  PRIOR FILING DATE: 2002-05-29
;  NUMBER OF SEQ ID NOS: 184
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
;   LENGTH: 98
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-379-392-21

  Query Match             100.0%;  Score 96;  DB 16;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-06;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         50 GINWNGGSTGYADSVKG 66


RESULT 13
US-10-447-331-6
; Sequence 6, Application US/10447331
; Publication No. US20030219434A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Paul J.
;  APPLICANT:  Ridgway, John B.
;  TITLE OF INVENTION: ANTIBODIES FOR CANCER THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS
;  FILE REFERENCE: GENENT.122A
;  CURRENT APPLICATION NUMBER: US/10/447,331
;  CURRENT FILING DATE:  2003-05-28
;  PRIOR APPLICATION NUMBER: US/09/515,825
;  PRIOR FILING DATE: 2000-02-29
;  PRIOR APPLICATION NUMBER: 60/122262
;  PRIOR FILING DATE: 1999-03-01
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 122
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-447-331-6

  Query Match             100.0%;  Score 96;  DB 15;  Length 122;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-06;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         50 GINWNGGSTGYADSVKG 66


RESULT 14
US-09-880-748-937
; Sequence 937, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Antibodies that Immunospecifically Bind BLyS
;  FILE REFERENCE: PF523
;  CURRENT APPLICATION NUMBER: US/09/880,748
;  CURRENT FILING DATE:  2001-06-15
;  PRIOR APPLICATION NUMBER: 60/212,210
;  PRIOR FILING DATE: 2000-06-15
;  PRIOR APPLICATION NUMBER: 60/240,816
;  PRIOR FILING DATE: 2000-10-17
;  PRIOR APPLICATION NUMBER: 60/276,248
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/277,379
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 60/293,499
;  PRIOR FILING DATE: 2001-05-25
;  NUMBER OF SEQ ID NOS: 3239
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 937
;   LENGTH: 239
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-880-748-937

  Query Match             100.0%;  Score 96;  DB 10;  Length 239;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-06;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         50 GINWNGGSTGYADSVKG 66


RESULT 15
US-09-880-748-2015
; Sequence 2015, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Antibodies that Immunospecifically Bind BLyS
;  FILE REFERENCE: PF523
;  CURRENT APPLICATION NUMBER: US/09/880,748
;  CURRENT FILING DATE:  2001-06-15
;  PRIOR APPLICATION NUMBER: 60/212,210
;  PRIOR FILING DATE: 2000-06-15
;  PRIOR APPLICATION NUMBER: 60/240,816
;  PRIOR FILING DATE: 2000-10-17
;  PRIOR APPLICATION NUMBER: 60/276,248
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/277,379
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 60/293,499
;  PRIOR FILING DATE: 2001-05-25
;  NUMBER OF SEQ ID NOS: 3239
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2015
;   LENGTH: 239
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-880-748-2015

  Query Match             100.0%;  Score 96;  DB 10;  Length 239;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-06;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         50 GINWNGGSTGYADSVKG 66


Search completed: September 18, 2004, 04:20:42
Job time : 675.143 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 03:33:04 ; Search time 34.6071 Seconds
(without alignments)
47.252 Million cell updates/sec

Title: US-10-029-926B-115
Perfect score: 96
Sequence: 1 GINWNGGSTGYADSVKG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	96	100.0	98	2	S26928	Ig heavy chain V r
2	87	90.6	112	2	PH1654	Ig heavy chain V r
3	79	82.3	98	2	S26927	Ig heavy chain V r
4	79	82.3	121	2	S31104	Ig heavy chain (su
5	79	82.3	128	2	S31595	Ig heavy chain V r
6	76	79.2	123	2	S30532	Ig heavy chain V r
7	74	77.1	121	2	S31118	Ig heavy chain - h
8	73	76.0	100	2	S69896	Ig heavy chain V r
9	71	74.0	98	2	S26929	Ig heavy chain V r
10	71	74.0	98	2	S26934	Ig heavy chain V r
11	71	74.0	120	2	S44111	Ig heavy chain V-D
12	71	74.0	191	2	JL0048	Ig heavy chain V r
13	69	71.9	120	1	G1HUDB	Ig heavy chain V-I
14	68	70.8	145	2	S11239	Ig heavy chain V r
15	65	67.7	120	2	S36273	Ig heavy chain V r
16	65	67.7	123	2	PC4281	anti-SS-A/Ro 60K p
17	64	66.7	90	2	S24248	Ig heavy chain V r
18	64	66.7	120	2	S36278	Ig heavy chain V r
19	62	64.6	96	2	S20781	Ig heavy chain V r
20	61.5	64.1	97	2	S26935	Ig heavy chain V r
21	61.5	64.1	97	2	S46462	Ig heavy chain V r
22	61.5	64.1	116	2	B28966	Ig heavy chain pre
23	61.5	64.1	116	2	S31110	Ig heavy chain - h
24	61.5	64.1	146	2	I47184	Ig heavy chain var
25	61	63.5	98	2	S26889	Ig heavy chain V r
26	61	63.5	98	2	S54856	Ig heavy chain V r
27	61	63.5	99	2	S24259	Ig heavy chain V r
28	61	63.5	100	2	S24258	Ig heavy chain V r
29	61	63.5	101	2	S24257	Ig heavy chain V r
30	61	63.5	102	2	S24260	Ig heavy chain V r
31	61	63.5	104	2	S24255	Ig heavy chain V r
32	61	63.5	105	2	S24249	Ig heavy chain V r
33	61	63.5	106	2	S24256	Ig heavy chain V r
34	61	63.5	108	2	PH1648	Ig heavy chain V r
35	61	63.5	109	2	PH1649	Ig heavy chain V r
36	61	63.5	109	2	S24254	Ig heavy chain V r
37	61	63.5	109	2	S24253	Ig heavy chain V r
38	61	63.5	110	2	S24250	Ig heavy chain V r
39	61	63.5	112	2	PH1647	Ig heavy chain V r
40	61	63.5	113	2	S24247	Ig heavy chain V r
41	61	63.5	115	2	S09382	Ig heavy chain - c
42	61	63.5	117	2	A45953	Ig heavy chain pre
43	61	63.5	118	2	S31121	Ig heavy chain - h
44	61	63.5	119	2	C36005	Ig heavy chain V r
45	61	63.5	119	2	D36005	Ig heavy chain V r

ALIGNMENTS

RESULT 1
S26928
Ig heavy chain V region (DP-32) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C;Accession: S26928
R;Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A;Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A;Reference number: S26885; MUID:93021117; PMID:1404388
A;Accession: S26928
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-98 <TOM>
A;Cross-references: EMBL:Z12334; NID:g32887; PIDN:CAA78204.1; PID:g32888
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 96; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GINWNGGSTGYADSVKG 17
              |||||||||||||||||
Db         50 GINWNGGSTGYADSVKG 66
```

RESULT 2
PH1654
Ig heavy chain V region (clone 6H9) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996
C;Accession: PH1654
R;Hillson, J.L.; Karr, N.S.; Oppliger, I.R.; Mannik, M.; Sasso, E.H.
J. Exp. Med. 178, 331-336, 1993
A;Title: The structural basis of germline-encoded VH3 immunoglobulin binding to staphyl
A;Reference number: PH1642; MUID:93301610; PMID:8315388
A;Accession: PH1654
A;Molecule type: mRNA
A;Residues: 1-112 <HIL>
A;Experimental source: B cell
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;7-90/Domain: immunoglobulin homology <IMM>

Query Match 90.6%; Score 87; DB 2; Length 112;
Best Local Similarity 94.1%; Pred. No. 5.6e-06;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GINWNGGSTGYADSVKG 17
              |||| ||||||||||||
```

```
Db         42 GINWIGGSTGYADSVKG 58


RESULT 3
S26927
Ig heavy chain V region (DP-31) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C;Accession: S26927
R;Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A;Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V
A;Reference number: S26885; MUID:93021117; PMID:1404388
A;Accession: S26927
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-98 <TOM>
A;Cross-references: EMBL:Z12333; NID:g32885; PIDN:CAA78203.1; PID:g32886
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

  Query Match             82.3%;  Score 79;  DB 2;  Length 98;
  Best Local Similarity   82.4%;  Pred. No. 7.8e-05;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||:|| ||`||||||||
Db         50 GISWNSGSIGYADSVKG 66


RESULT 4
S31104
Ig heavy chain (subclass IgM) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 23-Jul-1999
C;Accession: S31104
R;Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurman
Eur. J. Immunol. 22, 247-251, 1992
A;Title: Restricted utilization of germ-line V(H)3 genes and short diverse third complem
A;Reference number: S31104; MUID:92111633; PMID:1730252
A;Accession: S31104
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-121 <RAA>
A;Cross-references: EMBL:X63080; NID:g32646; PIDN:CAA44802.1; PID:g32647
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

  Query Match             82.3%;  Score 79;  DB 2;  Length 121;
  Best Local Similarity   82.4%;  Pred. No. 9.6e-05;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||:|| || ||||||||
Db         50 GISWNSGSIGYADSVKG 66


RESULT 5
S31595
Ig heavy chain V region - human
C;Species: Homo sapiens (man)
C;Date: 03-Mar-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C;Accession: S31595
R;Cuisinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnelle, C.
submitted to the EMBL Data Library, June 1992
A;Description: Mechanisms that generate human immunoglobulin diversity operate from the
A;Reference number: S31585
A;Accession: S31595
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-128 <CUI>
A;Cross-references: EMBL:Z14171; NID:g31007; PIDN:CAA78540.1; PID:g31008
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;23-106/Domain: immunoglobulin homology <IMM>

  Query Match             82.3%;  Score 79;  DB 2;  Length 128;
  Best Local Similarity   82.4%;  Pred. No. 0.0001;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||:|| || ||||||||
Db         58 GISWNSGSIGYADSVKG 74


RESULT 6
S30532
Ig heavy chain V region - human
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Aug-1996
C;Accession: S30532
R;Mariette, X.
submitted to the EMBL Data Library, October 1992
A;Reference number: S30520
A;Accession: S30532
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-123 <MAR>
A;Cross-references: EMBL:Z18318
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

  Query Match             79.2%;  Score 76;  DB 2;  Length 123;
  Best Local Similarity   76.5%;  Pred. No. 0.00027;
  Matches   13;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||:|| |: ||||||||
Db         50 GISWNSGTLGYADSVKG 66


RESULT 7
S31118
Ig heavy chain - human
C;Species: Homo sapiens (man)
C;Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 17-Mar-1999
C;Accession: S31118
R;Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurman
Eur. J. Immunol. 22, 247-251, 1992
A;Title: Restricted utilization of germ-line V(H)3 genes and short diverse third complem
A;Reference number: S31104; MUID:92111633; PMID:1730252
A;Accession: S31118
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-121 <RAA>
A;Cross-references: EMBL:X62969
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

  Query Match             77.1%;  Score 74;  DB 2;  Length 121;
  Best Local Similarity   76.5%;  Pred. No. 0.00054;
  Matches   13;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||:|| || || |||||
Db         50 GISWNSGSIGYRDSVKG 66


RESULT 8
S69896
```

Ig heavy chain V region (clone RFKL5H), rheumatoid factor - human
C;Species: Homo sapiens (man)
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 23-Jul-1999
C;Accession: S69896
R;Randen, I.; Pascual, V.; Victor, K.; Thompson, K.M.; Forre, O.; Capra, D.J.; Natvig, J
Eur. J. Immunol. 23, 1220-1225, 1993
A;Title: Synovial IgG rheumatoid factors show evidence of an antigen-driven immune respo
A;Reference number: S69896; MUID:93272805; PMID:8500520
A;Accession: S69896
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 1-100 <RAN>
A;Cross-references: EMBL:X73605; NID:g509797; PIDN:CAA51998.1; PID:g509798
C;Superfamily: immunoglobulin V region; immunoglobulin homology
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 76.0%; Score 73; DB 2; Length 100;
Best Local Similarity 76.5%; Pred. No. 0.00063;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
|| || | ||||||||
Db 50 GITWNSGRIGYADSVKG 66

RESULT 9
S26929
Ig heavy chain V region (DP-33) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C;Accession: S26929
R;Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A;Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V
A;Reference number: S26885; MUID:93021117; PMID:1404388
A;Accession: S26929
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-98 <TOM>
A;Cross-references: EMBL:Z12335; NID:g32889; PIDN:CAA78205.1; PID:g32890
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 74.0%; Score 71; DB 2; Length 98;
Best Local Similarity 81.2%; Pred. No. 0.0012;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 INWNGGSTGYADSVKG 17
|:|:|||| |||||||
Db 51 ISWDGGSTYYADSVKG 66

RESULT 10
S26934
Ig heavy chain V region (DP-59) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C;Accession: S26934
R;Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A;Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V
A;Reference number: S26885; MUID:93021117; PMID:1404388
A;Accession: S26934
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-98 <TOM>
A;Cross-references: EMBL:Z12359; NID:g32937; PIDN:CAA78229.1; PID:g32938
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 74.0%; Score 71; DB 2; Length 98;
Best Local Similarity 70.6%; Pred. No. 0.0012;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
|::||| | |||||||
Db 50 GVSWNGSRTHYADSVKG 66

RESULT 11
S44111
Ig heavy chain V-D-J region - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-May-2001
C;Accession: S44111
R;Hawkins, R.E.; Zhu, D.; Ovecka, M.; Winter, G.; Hamblin, T.J.; Stevenson, F.K.
submitted to the EMBL Data Library, March 1994
A;Description: Idiotypic vaccination against human B-cell lymphoma: rescue of variable
A;Reference number: S44105
A;Accession: S44111
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-120 <HAW>
A;Cross-references: EMBL:Z31387; NID:g472965; PIDN:CAA83262.1; PID:g940522
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 74.0%; Score 71; DB 2; Length 120;
Best Local Similarity 81.2%; Pred. No. 0.0015;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 INWNGGSTGYADSVKG 17
|:|:|||| |||||||
Db 51 ISWDGGSTYYADSVKG 66

RESULT 12
JL0048
Ig heavy chain V region precursor (clone cR19) - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 07-May-1999
C;Accession: JL0048; S66433
R;Baer, R.; Forster, A.; Lavenir, I.; Rabbitts, T.H.
J. Exp. Med. 167, 2011-2016, 1988
A;Title: Immunoglobulin V-H genes are transcribed by T cells in association with a new
A;Reference number: JL0047; MUID:88258392; PMID:3133445
A;Accession: JL0048
A;Molecule type: mRNA
A;Residues: 1-191 <BEA>
A;Note: the authors translated the reading frame which extends to the stop codon; the s
A;Note: this sequence belongs to the VH III subgroup
R;Mueller-Newen, G.; Koehne, C.; Keul, R.; Hemmann, U.; Mueller-Esterl, W.; Wijdenes, J
Eur. J. Biochem. 236, 837-842, 1996
A;Title: Purification and characterization of the soluble interleukin-6 receptor from h
A;Reference number: S66433; MUID:96270733; PMID:8665902
A;Accession: S66433
A;Molecule type: protein
A;Residues: 9-18 <MUE>
C;Superfamily: immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;23-106/Domain: immunoglobulin homology <IMM>

Query Match 74.0%; Score 71; DB 2; Length 191;
Best Local Similarity 70.6%; Pred. No. 0.0024;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
|::||| | |||||||
Db 58 GVSWNGSRTHYADSVKG 74

RESULT 13
G1HUDB

```
Ig heavy chain V-III region (Dob) - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1979 #sequence_revision 31-Dec-1979 #text_change 02-Sep-1997
C;Accession: A90431; A02065
R;Steiner, L.A.; Garcia Pardo, A.; Margolies, M.N.
Biochemistry 18, 4068-4080, 1979
A;Title: Amino acid sequence of the heavy-chain variable region of the crystallizable hu
A;Reference number: A90431; MUID:80020921; PMID:114209
A;Accession: A90431
A;Molecule type: protein
A;Residues: 1-120 <STE>
R;Steiner, L.A.; Lopes, A.D.
Biochemistry 18, 4054-4067, 1979
A;Title: The crystallizable human myeloma protein Dob has a hinge-region deletion.
A;Reference number: A90430; MUID:80020920; PMID:114208
A;Contents: annotation
A;Note: this gamma-1 myeloma protein has a deletion in the hinge region; there are no li
C;Genetics:
A;Gene: GDB:IGHV@
A;Cross-references: GDB:128528; OMIM:147070
A;Map position: 14q32.33-14q32.33
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>
F;22-96/Disulfide bonds: #status predicted

  Query Match             71.9%;  Score 69;  DB 1;  Length 120;
  Best Local Similarity   81.2%;  Pred. No. 0.003;
  Matches   13;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              | |||||  |||||||
Db         51 ITWNGGSVLYADSVKG 66


RESULT 14
S11239
Ig heavy chain V region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C;Accession: S11239
R;Felgenhauer, M.; Kohl, J.; Rueker, F.
Nucleic Acids Res. 18, 4927, 1990
A;Title: Nucleotide sequences of the cDNAs encoding the V-regions of H- and L-chains of
A;Reference number: S11239; MUID:90370490; PMID:1697678
A;Accession: S11239
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-145 <FEL>
A;Cross-references: EMBL:X53613; NID:g23865; PIDN:CAA37675.1; PID:g762936
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;34-117/Domain: immunoglobulin homology <IMM>

  Query Match             70.8%;  Score 68;  DB 2;  Length 145;
  Best Local Similarity   70.6%;  Pred. No. 0.0051;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||:|:  | ||||||||
Db         69 GISWDSSSIGYADSVKG 85


RESULT 15
S36273
Ig heavy chain V region (clone alpha-THY-32) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999
C;Accession: S36273
R;Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.
EMBO J. 12, 725-734, 1993
A;Title: Human anti-self antibodies with high specificity from phage display libraries.
A;Reference number: S36256; MUID:93178448; PMID:7679990
A;Accession: S36273
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-120 <GRI>
A;Cross-references: EMBL:Z18834; NID:g33116; PIDN:CAA79286.1; PID:g939896
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

  Query Match             67.7%;  Score 65;  DB 2;  Length 120;
  Best Local Similarity   76.5%;  Pred. No. 0.012;
  Matches   13;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||: | || ||||||||
Db         50 GISSNSGSIGYADSVKG 66


Search completed: September 18, 2004, 03:47:33
Job time : 35.6071 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 02:29:47 ; Search time 24.8929 Seconds
(without alignments)
35.560 Million cell updates/sec

Title: US-10-029-926B-115
Perfect score: 96
Sequence: 1 GINWNGGSTGYADSVKG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	69	71.9	120	1	HV3U_HUMAN	P01782 homo sapien
2	61.5	64.1	116	1	HV05_CARAU	P19181 carassius a
3	59	61.5	117	1	HV53_MOUSE	P18524 mus musculu
4	57	59.4	117	1	HV3C_HUMAN	P01764 homo sapien
5	53	55.2	115	1	HV3F_HUMAN	P01767 homo sapien
6	51	53.1	121	1	HV3J_HUMAN	P01771 homo sapien
7	50.5	52.6	97	1	HV56_MOUSE	P18527 mus musculu
8	48.5	50.5	117	1	HV02_CANFA	P01785 canis famil
9	48	50.0	117	1	HV55_MOUSE	P18526 mus musculu
10	48	50.0	119	1	HV3L_HUMAN	P01773 homo sapien
11	47	49.0	117	1	HV54_MOUSE	P18525 mus musculu
12	46.5	48.4	135	1	HV02_XENLA	P20957 xenopus lae
13	46	47.9	114	1	HV3B_HUMAN	P01763 homo sapien
14	46	47.9	115	1	HV3D_HUMAN	P01765 homo sapien
15	46	47.9	117	1	HV03_CARAU	P19180 carassius a
16	46	47.9	119	1	HV3M_HUMAN	P01774 homo sapien
17	46	47.9	119	1	HV3N_HUMAN	P01775 homo sapien
18	46	47.9	122	1	HV3A_HUMAN	P01762 homo sapien
19	46	47.9	126	1	HV3K_HUMAN	P01772 homo sapien
20	46	47.9	428	1	SP7_MOUSE	Q8vi67 mus musculu
21	46	47.9	431	1	SP7_HUMAN	Q8tdd2 homo sapien
22	46	47.9	559	1	PMGI_MAIZE	P30792 zea mays (m
23	45.5	47.4	111	1	HV35_MOUSE	P01804 mus musculu
24	45	46.9	116	1	HV3T_HUMAN	P01781 homo sapien
25	45	46.9	116	1	HV45_MOUSE	P01821 mus musculu
26	45	46.9	117	1	HV1B_HUMAN	P01743 homo sapien
27	45	46.9	121	1	HV01_MOUSE	P01745 mus musculu
28	45	46.9	122	1	HV3G_HUMAN	P01768 homo sapien
29	45	46.9	954	1	XYNA_RUMFL	P29126 ruminococcu
30	44.5	46.4	120	1	HV3E_HUMAN	P01766 homo sapien
31	44	45.8	115	1	HV32_MOUSE	P01801 mus musculu
32	44	45.8	725	1	AREA_PENCH	Q01582 penicillium
33	44	45.8	865	1	NRFA_PENUR	Q92269 penicillium
34	43	44.8	136	1	HV16_MOUSE	P01783 mus musculu
35	43	44.8	151	1	YF84_METJA	Q58979 methanococc
36	43	44.8	327	1	K6P1_BACTN	Q8a8r5 bacteroides
37	43	44.8	527	1	PPCK_OCEIH	Q8ep04 oceanobacil
38	43	44.8	649	1	PRTS_MACMU	Q28520 macaca mula
39	43	44.8	676	1	PRTS_HUMAN	P07225 homo sapien
40	43	44.8	678	1	GSPD_AERSA	P45778 aeromonas s
41	43	44.8	731	1	HUTU_CAEEL	Q9nae2 caenorhabdi
42	43	44.8	835	1	FASD_ECOLI	P46000 escherichia
43	43	44.8	1019	1	CA16_CHICK	P20785 gallus gall
44	43	44.8	1633	1	YP74_CAEEL	Q09221 caenorhabdi
45	42.5	44.3	264	1	EXB5_ARATH	Q9m203 arabidopsis

ALIGNMENTS

RESULT 1
HV3U_HUMAN
ID HV3U_HUMAN STANDARD; PRT; 120 AA.
AC P01782;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-III region DOB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=80020921; PubMed=114209;
RA Steiner L.A., Garcia Pardo A., Margolies M.N.;
RT "Amino acid sequence of the heavy-chain variable region of the
RT crystallizable human myeloma protein Dob.";
RL Biochemistry 18:4068-4080(1979).
RN [2]
RP CRYSTALLIZATION.
RX MEDLINE=80020920; PubMed=114208;
RA Steiner L.A., Lopes A.D.;
RT "The crystallizable human myeloma protein Dob has a hinge-region
RT deletion.";
RL Biochemistry 18:4054-4067(1979).
CC -!- MISCELLANEOUS: THIS GAMMA-1 MYELOMA PROTEIN HAS A DELETION IN THE
CC HINGE REGION. THERE ARE NO LIGHT-HEAVY OR INTER-HEAVY CHAIN
CC DISULFIDE BONDS.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A90431; G1HUDB.
DR HSSP; P01772; 2FB4.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 112 IG-LIKE.
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13440 MW; 880DDE307C4B2627 CRC64;

Query Match 71.9%; Score 69; DB 1; Length 120;
Best Local Similarity 81.2%; Pred. No. 0.00061;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 INWNGGSTGYADSVKG 17
              | |||||  |||||||
Db         51 ITWNGGSVLYADSVKG 66
```

RESULT 2
HV05_CARAU

```
ID   HV05_CARAU     STANDARD;      PRT;   116 AA.
AC   P19181;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig heavy chain V region 5A precursor.
OS   Carassius auratus (Goldfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Carassius.
OX   NCBI_TaxID=7957;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88144476; PubMed=3125551;
RA   Wilson M.R., Middleton D., Warr G.W.;
RT   "Immunoglobulin heavy chain variable region gene evolution: structure
RT   and family relationships of two genes and a pseudogene in a teleost
RT   fish.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).
DR   HSSP; P01772; 2FB4.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region; Signal.
FT   SIGNAL        1     19
FT   CHAIN         20    116       IG HEAVY CHAIN V REGION 5A.
FT   DOMAIN        20     49       FRAMEWORK-1.
FT   DOMAIN        50     54       COMPLEMENTARITY-DETERMINING-1.
FT   DOMAIN        55     68       FRAMEWORK-2.
FT   DOMAIN        69     84       COMPLEMENTARITY-DETERMINING-2.
FT   DOMAIN        85    116       FRAMEWORK-3.
FT   DISULFID      41    114       BY SIMILARITY.
FT   NON_TER      116    116
SQ   SEQUENCE   116 AA;  12808 MW;  9C2279E2DF199B12 CRC64;

  Query Match             64.1%;  Score 61.5;  DB 1;  Length 116;
  Best Local Similarity   59.1%;  Pred. No. 0.0088;
  Matches   13;  Conservative    2;  Mismatches    2;  Indels    5;  Gaps    1;

Qy          1 GINW-----NGGSTGYADSVKG 17
              |: |     :|||| |||||||
Db         63 GLEWVSVIYSGGSTYYADSVKG 84


RESULT 3
HV53_MOUSE
ID   HV53_MOUSE     STANDARD;      PRT;   117 AA.
AC   P18524;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig heavy chain V region RF precursor.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/cJ;
RX   MEDLINE=89279149; PubMed=2499654;
RA   Levy N.S., Malipiero U.V., Lebecque S.G., Gearhart P.J.;
RT   "Early onset of somatic mutation in immunoglobulin VH genes during
RT   the primary immune response.";
RL   J. Exp. Med. 169:2007-2019(1989).
CC   -!- MISCELLANEOUS: THIS SEQUENCE BELONGS TO THE VH7183 SUBFAMILY.
DR   PIR; JT0503; HVMSRF.
DR   HSSP; P01810; 2FBJ.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region; Hybridoma; Signal.
FT   SIGNAL        1     19
FT   CHAIN         20    117       IG HEAVY CHAIN V REGION RF.
FT   DOMAIN        20     49       FRAMEWORK-1.
FT   DOMAIN        50     54       COMPLEMENTARITY-DETERMINING-1.
FT   DOMAIN        55     68       FRAMEWORK-2.
FT   DOMAIN        69     85       COMPLEMENTARITY-DETERMINING-2.
FT   DOMAIN        86    117       FRAMEWORK-3.
FT   DISULFID      41    115       BY SIMILARITY.
FT   NON_TER      117    117
SQ   SEQUENCE   117 AA;  12866 MW;  2CE3295F390F725B CRC64;

  Query Match             61.5%;  Score 59;  DB 1;  Length 117;
  Best Local Similarity   75.0%;  Pred. No. 0.022;
  Matches   12;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              || ||||| | |:|||
Db         70 INSNGGSTYYPDTVKG 85


RESULT 4
HV3C_HUMAN
ID   HV3C_HUMAN     STANDARD;      PRT;   117 AA.
AC   P01764;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ig heavy chain V-III region VH26 precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81101090; PubMed=6450418;
RA   Matthyssens G., Rabbitts T.H.;
RT   "Structure and multiplicity of genes for the human immunoglobulin
RT   heavy chain variable region.";
RL   Proc. Natl. Acad. Sci. U.S.A. 77:6561-6565(1980).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J00236; AAA53516.1; -.
DR   EMBL; M35415; AAA58735.1; -.
DR   PIR; A02047; H3HU26.
DR   PDB; 1HOU; 23-DEC-99.
DR   Genew; HGNC:5545; IGHV@.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0003823; F:antigen binding; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region; Signal; 3D-structure.
FT   SIGNAL        1     19
FT   CHAIN         20    117       IG HEAVY CHAIN V-III REGION VH26.
FT   DOMAIN        20   >117       IG-LIKE.
FT   NON_TER      117    117
SQ   SEQUENCE   117 AA;  12582 MW;  E826733F1A3CB0F1 CRC64;

  Query Match             59.4%;  Score 57;  DB 1;  Length 117;
  Best Local Similarity   52.2%;  Pred. No. 0.045;
```

```
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    6;  Gaps    1;

Qy          1 GINW------NGGSTGYADSVKG 17
              |: |      :|||| | |||||
Db         63 GLEWVSAISGSGGGSTYYGDSVKG 85


RESULT 5
HV3F_HUMAN
ID   HV3F_HUMAN     STANDARD;      PRT;   115 AA.
AC   P01767;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ig heavy chain V-III region BUT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=78137069; PubMed=416441;
RA   Torano A., Putnam F.W.;
RT   "Complete amino acid sequence of the alpha 2 heavy chain of a human
RT   IgA2 immunoglobulin of the A2m (2) allotype.";
RL   Proc. Natl. Acad. Sci. U.S.A. 75:966-969(1978).
CC   -!- MISCELLANEOUS: THE SEQUENCE OF THE ALPHA-2, A2M(2) ALLOTYPE, C
CC       REGION OF THIS MYELOMA PROTEIN IS ALSO GIVEN.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR   PIR; A02050; A2HUBU.
DR   HSSP; P01789; 1MCP.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0003823; F:antigen binding; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region.
FT   DOMAIN        1    111       IG-LIKE.
FT   NON_TER     115    115
SQ   SEQUENCE   115 AA;  12379 MW;  208876A7DF52DCF4 CRC64;

  Query Match             55.2%;  Score 53;  DB 1;  Length 115;
  Best Local Similarity   71.4%;  Pred. No. 0.19;
  Matches   10;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 WNGGSTGYADSVKG 17
              : ||:| |||||||
Db         52 YRGGTTYYADSVKG 65


RESULT 6
HV3J_HUMAN
ID   HV3J_HUMAN     STANDARD;      PRT;   121 AA.
AC   P01771;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ig heavy chain V-III region HIL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=79124695; PubMed=420800;
RA   Chiu Y.-Y.H., Lopez de Castro J.A., Poljak R.J.;
RT   "Amino acid sequence of the VH region of human myeloma
RT   cryoimmunoglobulin IgG Hil.";
RL   Biochemistry 18:553-560(1979).
CC   -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
CC       PROTEIN.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR   PIR; A02054; G1HUHL.
DR   HSSP; P01772; 2FB4.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0003823; F:antigen binding; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region; Pyrrolidone carboxylic acid.
FT   DOMAIN        1    112       IG-LIKE.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   NON_TER     121    121
SQ   SEQUENCE   121 AA;  13566 MW;  480FC53610EF5DAB CRC64;

  Query Match             53.1%;  Score 51;  DB 1;  Length 121;
  Best Local Similarity   47.8%;  Pred. No. 0.41;
  Matches   11;  Conservative    1;  Mismatches    5;  Indels    6;  Gaps    1;

Qy          1 GINW------NGGSTGYADSVKG 17
              |: |      ||  | | |||||
Db         44 GLEWVAVIWYNGSRTYYGDSVKG 66


RESULT 7
HV56_MOUSE
ID   HV56_MOUSE     STANDARD;      PRT;    97 AA.
AC   P18527;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ig heavy chain V region 914.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/cJ;
RX   MEDLINE=89279149; PubMed=2499654;
RA   Levy N.S., Malipiero U.V., Lebecque S.G., Gearhart P.J.;
RT   "Early onset of somatic mutation in immunoglobulin VH genes during
RT   the primary immune response.";
RL   J. Exp. Med. 169:2007-2019(1989).
CC   -!- MISCELLANEOUS: THIS SEQUENCE BELONGS TO THE VH7183 SUBFAMILY.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR   PIR; JT0504; HVMS91.
DR   HSSP; P01810; 2FBJ.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region.
FT   DOMAIN        1    >97       IG-LIKE.
FT   NON_TER      97     97
SQ   SEQUENCE   97 AA;  10661 MW;  C23CB33FF55DA893 CRC64;

  Query Match             52.6%;  Score 50.5;  DB 1;  Length 97;
  Best Local Similarity   52.4%;  Pred. No. 0.39;
  Matches   11;  Conservative    2;  Mismatches    3;  Indels    5;  Gaps    1;

Qy          2 INW-----NGGSTGYADSVKG 17
              : |      :|||| | |||||
Db         45 LEWVASISSGGSTYYPDSVKG 65


RESULT 8
HV02_CANFA
ID   HV02_CANFA     STANDARD;      PRT;   117 AA.
```

```
AC   P01785;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ig heavy chain V region MOO.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE OF 1-112.
RX   MEDLINE=77242268; PubMed=407924;
RA   Wasserman R.L., Capra J.D.;
RT   "Primary structure of the variable regions of two canine
RT   immunoglobulin heavy chains.";
RL   Biochemistry 16:3160-3168(1977).
RN   [2]
RP   SEQUENCE OF 113-117.
RX   MEDLINE=80077682; PubMed=117299;
RA   McCumber L.J., Capra J.D.;
RT   "The complete amino-acid sequence of a canine mu chain.";
RL   Mol. Immunol. 16:565-570(1979).
CC   -!- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR   PIR; A90403; MHDGMO.
DR   HSSP; P01772; 2FB4.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region.
FT   DOMAIN        1    116       IG-LIKE.
FT   NON_TER     117    117
SQ   SEQUENCE   117 AA;  12703 MW;  FCE07309E0A84B35 CRC64;

  Query Match             50.5%;  Score 48.5;  DB 1;  Length 117;
  Best Local Similarity   45.5%;  Pred. No. 0.97;
  Matches   10;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 GINW-----NGGSTGYADSVKG 17
              |: |     : | | |||:|||
Db         44 GLQWVADISSSGQTYYADAVKG 65


RESULT 9
HV55_MOUSE
ID   HV55_MOUSE     STANDARD;      PRT;   117 AA.
AC   P18526;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig heavy chain V region 345 precursor.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/cJ;
RX   MEDLINE=89279149; PubMed=2499654;
RA   Levy N.S., Malipiero U.V., Lebecque S.G., Gearhart P.J.;
RT   "Early onset of somatic mutation in immunoglobulin VH genes during
RT   the primary immune response.";
RL   J. Exp. Med. 169:2007-2019(1989).
CC   -!- MISCELLANEOUS: THIS SEQUENCE BELONGS TO THE VH7183 SUBFAMILY.
DR   PIR; JT0502; HVMS34.
DR   HSSP; P01810; 2FBJ.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    117       IG HEAVY CHAIN V REGION 345.
FT   DOMAIN       20     49       FRAMEWORK-1.
FT   DOMAIN       50     54       COMPLEMENTARITY-DETERMINING-1.
FT   DOMAIN       55     68       FRAMEWORK-2.
FT   DOMAIN       69     85       COMPLEMENTARITY-DETERMINING-2.
FT   DOMAIN       86    117       FRAMEWORK-3.
FT   DISULFID     41    115       BY SIMILARITY.
FT   NON_TER     117    117
SQ   SEQUENCE   117 AA;  12902 MW;  49380E4627ACA99A CRC64;

  Query Match             50.0%;  Score 48;  DB 1;  Length 117;
  Best Local Similarity   62.5%;  Pred. No. 1.2;
  Matches   10;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              |:  |||| | |:|||
Db         70 ISSGGGSTYYPDTVKG 85


RESULT 10
HV3L_HUMAN
ID   HV3L_HUMAN     STANDARD;      PRT;   119 AA.
AC   P01773;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ig heavy chain V-III region BUR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE (MYELOMA PROTEIN BUR).
RX   MEDLINE=79151016; PubMed=107164;
RA   Putnam F.W., Liu Y.-S.V., Low T.L.K.;
RT   "Primary structure of a human IgA1 immunoglobulin. IV. Streptococcal
RT   IgA1 protease, digestion, Fab and Fc fragments, and the complete
RT   amino acid sequence of the alpha 1 heavy chain.";
RL   J. Biol. Chem. 254:2865-2874(1979).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR   PIR; A02056; A1HUBR.
DR   HSSP; P01772; 2FB4.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0003823; F:antigen binding; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region; Glycoprotein; Pyrrolidone carboxylic acid.
FT   DOMAIN        1    112       IG-LIKE.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   DISULFID     22     96
FT   CARBOHYD     28     28       N-LINKED (GLCNAC. . .).
FT   NON_TER     119    119
SQ   SEQUENCE   119 AA;  12981 MW;  12A709A75344D024 CRC64;

  Query Match             50.0%;  Score 48;  DB 1;  Length 119;
  Best Local Similarity   56.2%;  Pred. No. 1.2;
  Matches    9;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              |:: |  | |||||:|
Db         51 ISYGGSBTYYADSVRG 66


RESULT 11
HV54_MOUSE
ID   HV54_MOUSE     STANDARD;      PRT;   117 AA.
AC   P18525;
```

```
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig heavy chain V region 5-84 precursor.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/cJ;
RX   MEDLINE=89279149; PubMed=2499654;
RA   Levy N.S., Malipiero U.V., Lebecque S.G., Gearhart P.J.;
RT   "Early onset of somatic mutation in immunoglobulin VH genes during
RT   the primary immune response.";
RL   J. Exp. Med. 169:2007-2019(1989).
CC   -!- MISCELLANEOUS: THIS SEQUENCE BELONGS TO THE VH7183 SUBFAMILY.
DR   PIR; JT0505; HVMS84.
DR   HSSP; P01810; 2FBJ.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    117       IG HEAVY CHAIN V REGION 5-84.
FT   DOMAIN       20     49       FRAMEWORK-1.
FT   DOMAIN       50     54       COMPLEMENTARITY-DETERMINING-1.
FT   DOMAIN       55     68       FRAMEWORK-2.
FT   DOMAIN       69     85       COMPLEMENTARITY-DETERMINING-2.
FT   DOMAIN       86    117       FRAMEWORK-3.
FT   DISULFID     41    115       BY SIMILARITY.
FT   NON_TER     117    117
SQ   SEQUENCE   117 AA;  12872 MW;  234055CB6A469861 CRC64;

  Query Match             49.0%;  Score 47;  DB 1;  Length 117;
  Best Local Similarity   62.5%;  Pred. No. 1.7;
  Matches   10;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              |: |||| | |:|||
Db         70 ISNGGGSTYYPDTVKG 85


RESULT 12
HV02_XENLA
ID   HV02_XENLA     STANDARD;      PRT;   135 AA.
AC   P20957;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ig heavy chain V region XIG14 precursor (Fragment).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88176921; PubMed=2451244;
RA   Schwager J., Mikoryak C.A., Steiner L.A.;
RT   "Amino acid sequence of heavy chain from Xenopus laevis IgM deduced
RT   from cDNA sequence: implications for evolution of immunoglobulin
RT   domains.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:2245-2249(1988).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03632; AAA49791.1; -.
DR   PIR; B31933; B31933.
DR   HSSP; P01810; 2FBJ.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     18
FT   CHAIN        19    135       IG HEAVY CHAIN V REGION XIG14.
FT   DOMAIN       20    128       IG-LIKE.
FT   NON_TER     135    135
SQ   SEQUENCE   135 AA;  15080 MW;  EBC467105C00732E CRC64;

  Query Match             48.4%;  Score 46.5;  DB 1;  Length 135;
  Best Local Similarity   47.6%;  Pred. No. 2.3;
  Matches   10;  Conservative    2;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 GINW-----NGGSTGYADSVK 16
              |: |      ||||  |||:|
Db         61 GLEWIGVIATGGSTAIADSLK 81


RESULT 13
HV3B_HUMAN
ID   HV3B_HUMAN     STANDARD;      PRT;   114 AA.
AC   P01763;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ig heavy chain V-III region WEA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=83273707; PubMed=6410398;
RA   Goni F., Frangione B.;
RT   "Amino acid sequence of the Fv region of a human monoclonal IgM
RT   (protein WEA) with antibody activity against 3,4-pyruvylated
RT   galactose in Klebsiella polysaccharides K30 and K33.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:4837-4841(1983).
CC   -!- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM A MONOCLONAL ANTIBODY
CC       AGAINST 3,4-PYRUVYLATED GALACTOSE AND ISOLATED FROM A PATIENT WITH
CC       WALDENSTROM'S MACROGLOBULINEMIA.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR   PIR; A02046; M3HUWE.
DR   HSSP; P01772; 2FB4.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0003823; F:antigen binding; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region; Pyrrolidone carboxylic acid.
FT   DOMAIN        1    112       IG-LIKE.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   NON_TER     114    114
SQ   SEQUENCE   114 AA;  12256 MW;  D88294FB418A07B7 CRC64;

  Query Match             47.9%;  Score 46;  DB 1;  Length 114;
  Best Local Similarity   43.5%;  Pred. No. 2.3;
  Matches   10;  Conservative    3;  Mismatches    4;  Indels    6;  Gaps    1;

Qy          1 GINW------NGGSTGYADSVKG 17
```

```
              |: |      :| :  |||||||
Db         44 GLEWLSFIGGSGGSTIYYADSVKG 66


RESULT 14
HV3D_HUMAN
ID   HV3D_HUMAN     STANDARD;      PRT;   115 AA.
AC   P01765;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ig heavy chain V-III region TIL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=78005528; PubMed=409716;
RA   Wang A.-C., Wang I.Y., Fudenberg H.H.;
RT   "Immunoglobulin structure and genetics. Identity between variable
RT   regions of a mu and a gamma2 chain.";
RL   J. Biol. Chem. 252:7192-7199(1977).
CC   -!- MISCELLANEOUS: THE SEQUENCES OF THE V REGIONS OF THE HEAVY CHAINS
CC       OF IGM AND IGG2 ISOLATED FROM A SINGLE PATIENT WITH BICLONAL
CC       GAMMOPATHY ARE IDENTICAL. THEIR LIGHT CHAINS ARE APPARENTLY ALSO
CC       IDENTICAL.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR   PIR; A02048; H3HUTL.
DR   HSSP; P01772; 2FB4.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0003823; F:antigen binding; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region.
FT   DOMAIN        1    108       IG-LIKE.
FT   NON_TER     115    115
SQ   SEQUENCE   115 AA;  12356 MW;  4DCC67D179F62326 CRC64;

  Query Match             47.9%;  Score 46;  DB 1;  Length 115;
  Best Local Similarity   43.5%;  Pred. No. 2.4;
  Matches   10;  Conservative    2;  Mismatches    5;  Indels    6;  Gaps    1;

Qy          1 GINWNGG------STGYADSVKG 17
              |: | |      |  ||:||||
Db         44 GLZWVGAIZGLSVSZSYABSVKG 66


RESULT 15
HV03_CARAU
ID   HV03_CARAU     STANDARD;      PRT;   117 AA.
AC   P19180;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig heavy chain V region 3 precursor.
OS   Carassius auratus (Goldfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Carassius.
OX   NCBI_TaxID=7957;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88144476; PubMed=3125551;
RA   Wilson M.R., Middleton D., Warr G.W.;
RT   "Immunoglobulin heavy chain variable region gene evolution: structure
RT   and family relationships of two genes and a pseudogene in a teleost
RT   fish.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03616; AAA50807.1; -.
DR   PIR; A28966; A28966.
DR   HSSP; P01772; 2FB4.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin V region; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    117       IG HEAVY CHAIN V REGION 3.
FT   DOMAIN       20     49       FRAMEWORK-1.
FT   DOMAIN       50     54       COMPLEMENTARITY-DETERMINING-1.
FT   DOMAIN       55     68       FRAMEWORK-2.
FT   DOMAIN       69     85       COMPLEMENTARITY-DETERMINING-2.
FT   DOMAIN       86    117       FRAMEWORK-3.
FT   DISULFID     41    115       BY SIMILARITY.
FT   NON_TER     117    117
SQ   SEQUENCE   117 AA;  13220 MW;  512B625003FA6ECB CRC64;

  Query Match             47.9%;  Score 46;  DB 1;  Length 117;
  Best Local Similarity   39.1%;  Pred. No. 2.4;
  Matches    9;  Conservative    3;  Mismatches    5;  Indels    6;  Gaps    1;

Qy          1 GINW------NGGSTGYADSVKG 17
              |: |      :|    |||:|||
Db         63 GLEWVSRIYDDGSDVSYADTVKG 85


Search completed: September 18, 2004, 03:43:47
Job time : 25.8929 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 18, 2004, 03:24:30 ; Search time 88.6429 Seconds
(without alignments)
60.510 Million cell updates/sec

Title: US-10-029-926B-115
Perfect score: 96
Sequence: 1 GINWNGGSTGYADSVKG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	96	100.0	112	4	Q9HCC1	Q9hcc1 homo sapien
2	84	87.5	499	4	Q8N5K4	Q8n5k4 homo sapien
3	79	82.3	104	4	Q9UL87	Q9ul87 homo sapien
4	79	82.3	573	4	Q8WU38	Q8wu38 homo sapien
5	61	63.5	121	4	Q9UL71	Q9ul71 homo sapien
6	61	63.5	597	4	Q96BB9	Q96bb9 homo sapien
7	56	58.3	118	4	Q9UL72	Q9ul72 homo sapien
8	55	57.3	479	11	Q91WP5	Q91wp5 mus musculu
9	54	56.2	112	4	Q9UGP3	Q9ugp3 homo sapien
10	51	53.1	469	2	Q9FDT2	Q9fdt2 shewanella
11	51	53.1	838	16	Q7UUV9	Q7uuv9 rhodopirell
12	49	51.0	767	5	Q9U234	Q9u234 caenorhabdi
13	48	50.0	279	16	Q9WZJ8	Q9wzj8 thermotoga
14	48	50.0	298	10	Q9XGB4	Q9xgb4 trifolium r
15	48	50.0	416	4	Q9NPP6	Q9npp6 homo sapien
16	48	50.0	469	16	Q8PDZ8	Q8pdz8 xanthomonas
17	48	50.0	473	11	Q91Z05	Q91z05 mus musculu
18	48	50.0	654	5	Q86AX3	Q86ax3 dictyosteli
19	48	50.0	761	5	Q20898	Q20898 caenorhabdi
20	48	50.0	4342	16	Q9I157	Q9i157 pseudomonas
21	47	49.0	87	9	Q7Y3K4	Q7y3k4 enterobacte
22	47	49.0	116	4	Q9UL93	Q9ul93 homo sapien
23	47	49.0	469	16	Q87CN9	Q87cn9 xylella fas
24	47	49.0	515	5	Q16968	Q16968 aplysia cal
25	47	49.0	613	4	Q8WUK1	Q8wuk1 homo sapien
26	46	47.9	113	4	Q9UL90	Q9ul90 homo sapien
27	46	47.9	122	4	Q9UL84	Q9ul84 homo sapien
28	46	47.9	131	4	Q9UL88	Q9ul88 homo sapien
29	46	47.9	159	16	Q92CR6	Q92cr6 listeria in
30	46	47.9	159	16	Q8Y7X8	Q8y7x8 listeria mo
31	46	47.9	260	16	Q8NPS3	Q8nps3 corynebacte
32	46	47.9	413	4	Q7Z718	Q7z718 homo sapien
33	46	47.9	428	11	Q811U1	Q811u1 rattus norv
34	46	47.9	464	10	Q8S7G0	Q8s7g0 oryza sativ
35	46	47.9	469	16	Q8PQW5	Q8pqw5 xanthomonas
36	46	47.9	473	2	Q9RBK1	Q9rbk1 xanthomonas
37	46	47.9	493	4	Q8NCL6	Q8ncl6 homo sapien
38	46	47.9	627	16	Q8P8D7	Q8p8d7 xanthomonas
39	46	47.9	660	16	Q8PJW5	Q8pjw5 xanthomonas
40	45.5	47.4	666	5	Q9VA77	Q9va77 drosophila
41	45	46.9	124	4	Q9UL92	Q9ul92 homo sapien
42	45	46.9	125	4	Q9UL95	Q9ul95 homo sapien
43	45	46.9	308	2	Q9S0K5	Q9s0k5 shewanella
44	45	46.9	376	15	Q7ZP49	Q7zp49 human immun
45	45	46.9	393	16	Q8YSL5	Q8ysl5 anabaena sp

ALIGNMENTS

RESULT 1
Q9HCC1
ID Q9HCC1 PRELIMINARY; PRT; 112 AA.
AC Q9HCC1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Single chain Fv (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kikuchi M., Takeda C., Tsujimoto Y., Asada S., Nagata K.;
RT "An antibody fragment2A3 specific for native lysozyme :Isolaion from a
RT human synthetic phage display library and characterization.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB049915; BAB16829.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1 1
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12243 MW; 24F1A45EC3B84788 CRC64;

Query Match 100.0%; Score 96; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 6e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GINWNGGSTGYADSVKG 17
|||||||||||||||||
Db 50 GINWNGGSTGYADSVKG 66

RESULT 2

```
Q8N5K4
ID   Q8N5K4      PRELIMINARY;      PRT;   499 AA.
AC   Q8N5K4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RA   Strausberg R.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC032249; AAH32249.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00409; IG; 4.
DR   SMART; SM00407; IGc1; 2.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   499 AA;  53376 MW;  93A5C89582054F32 CRC64;

  Query Match             87.5%;  Score 84;  DB 4;  Length 499;
  Best Local Similarity   93.8%;  Pred. No. 0.00022;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              ||||||| |||||||
Db         70 INWNGGSTNYADSVKG 85


RESULT 3
Q9UL87
ID   Q9UL87      PRELIMINARY;      PRT;   104 AA.
AC   Q9UL87;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Myosin-reactive immunoglobulin heavy chain variable region
DE   (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98277139; PubMed=9614934;
RA   Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA   Young D.C.;
RT   "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT   fetus.";
RL   Clin. Immunol. Immunopathol. 87:184-192(1998).
DR   EMBL; AF035027; AAD56263.1; -.
DR   HSSP; P01810; 2FBJ.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
FT   NON_TER       1      1
FT   NON_TER     104    104
SQ   SEQUENCE   104 AA;  11598 MW;  611D1A3F40E96E7B CRC64;

  Query Match             82.3%;  Score 79;  DB 4;  Length 104;
  Best Local Similarity   82.4%;  Pred. No. 0.00021;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||:|| || ||||||||
Db         29 GISWNSGSIGYADSVKG 45


RESULT 4
Q8WU38
ID   Q8WU38      PRELIMINARY;      PRT;   573 AA.
AC   Q8WU38;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Tonsil;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC021276; AAH21276.1; -.
DR   PIR; S21205; S21205.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; 2.
KW   Hypothetical protein.
SQ   SEQUENCE   573 AA;  62967 MW;  FD072344033AC530 CRC64;

  Query Match             82.3%;  Score 79;  DB 4;  Length 573;
  Best Local Similarity   82.4%;  Pred. No. 0.0015;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              ||:|| || ||||||||
Db         69 GISWNSGSIGYADSVKG 85


RESULT 5
Q9UL71
ID   Q9UL71      PRELIMINARY;      PRT;   121 AA.
AC   Q9UL71;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Myosin-reactive immunoglobulin heavy chain variable region
DE   (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98277139; PubMed=9614934;
RA   Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA   Young D.C.;
RT   "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT   fetus.";
RL   Clin. Immunol. Immunopathol. 87:184-192(1998).
DR   EMBL; AF035043; AAD56279.1; -.
DR   HSSP; P01772; 2FB4.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
```

```
DR   PROSITE; PS50835; IG_LIKE; 1.
FT   NON_TER       1      1
FT   NON_TER     121    121
SQ   SEQUENCE   121 AA;  13154 MW;  2F045CCFA5D50736 CRC64;

  Query Match             63.5%;  Score 61;  DB 4;  Length 121;
  Best Local Similarity   56.5%;  Pred. No. 0.14;
  Matches   13;  Conservative    2;  Mismatches    2;  Indels    6;  Gaps    1;

Qy          1 GINW------NGGSTGYADSVKG 17
              |: |      :|||| |||||||
Db         44 GLEWVSLISGDGGSTYYADSVKG 66


RESULT 6
Q96BB9
ID   Q96BB9      PRELIMINARY;      PRT;   597 AA.
AC   Q96BB9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-cell;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC015760; AAH15760.1; -.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 5.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 5.
DR   PROSITE; PS00290; IG_MHC; 3.
KW   Hypothetical protein.
SQ   SEQUENCE   597 AA;  65039 MW;  4FCA3AD8ECE263D9 CRC64;

  Query Match             63.5%;  Score 61;  DB 4;  Length 597;
  Best Local Similarity   56.5%;  Pred. No. 0.86;
  Matches   13;  Conservative    2;  Mismatches    2;  Indels    6;  Gaps    1;

Qy          1 GINW------NGGSTGYADSVKG 17
              |: |      :|||| |||||||
Db         63 GLEWVSAISGSGGSTYYADSVKG 85


RESULT 7
Q9UL72
ID   Q9UL72      PRELIMINARY;      PRT;   118 AA.
AC   Q9UL72;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Myosin-reactive immunoglobulin heavy chain variable region
DE   (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98277139; PubMed=9614934;
RA   Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA   Young D.C.;
RT   "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT   fetus.";
RL   Clin. Immunol. Immunopathol. 87:184-192(1998).
DR   EMBL; AF035042; AAD56278.1; -.
DR   PIR; S21205; S21205.
DR   HSSP; P01772; 2FB4.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
FT   NON_TER       1      1
FT   NON_TER     118    118
SQ   SEQUENCE   118 AA;  12872 MW;  B4D1A5944B2D5CCA CRC64;

  Query Match             58.3%;  Score 56;  DB 4;  Length 118;
  Best Local Similarity   62.5%;  Pred. No. 0.78;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              : ::|||: |||||||
Db         50 VTYSGGGSSYYADSVKG 65


RESULT 8
Q91WP5
ID   Q91WP5      PRELIMINARY;      PRT;   479 AA.
AC   Q91WP5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Colon;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC013656; AAH13656.1; -.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; 2.
KW   Hypothetical protein.
SQ   SEQUENCE   479 AA;  51603 MW;  ECB2D0877748584F CRC64;

  Query Match             57.3%;  Score 55;  DB 11;  Length 479;
  Best Local Similarity   62.5%;  Pred. No. 5.5;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              || |||:| |:|::||
Db         70 INSNGGNTYYSDTMKG 85


RESULT 9
Q9UGP3
ID   Q9UGP3      PRELIMINARY;      PRT;   112 AA.
AC   Q9UGP3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Immunoblobulin heavy chain (Fragment).
GN   IGH.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zafiropoulos A., Kandilogiannaki M., Dahlenborg C., Borrebaeck C.A.K.,
```

```
RA   Krambovitis E.;
RT   "Induction of somatic mutations in human B cells by in vitro
RT   immunization.";
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ132560; CAB65078.1; -.
DR   HSSP; P01772; 2FB4.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
FT   NON_TER       1      1
FT   NON_TER     112    112
SQ   SEQUENCE   112 AA;  12226 MW;  0A1F17082C7A8CE3 CRC64;

  Query Match             56.2%;  Score 54;  DB 4;  Length 112;
  Best Local Similarity   68.8%;  Pred. No. 1.5;
  Matches   11;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVKG 17
              ||    |:| |||||||
Db         34 INATSGTTDYADSVKG 49


RESULT 10
Q9FDT2
ID   Q9FDT2      PRELIMINARY;      PRT;   469 AA.
AC   Q9FDT2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia
DE   ligase).
GN   GLNA.
OS   Shewanella benthica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC   Alteromonadaceae; Shewanella.
OX   NCBI_TaxID=43661;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DB6705;
RA   Kato C.;
RT   "Glutamine synthetase from Shewanella benthica strain DB6705.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + PHOSPHATE +
CC       L-GLUTAMINE.
CC   -!- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO
CC       HEXAGONS (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
DR   EMBL; AB038390; BAB11930.1; -.
DR   HSSP; P06201; 1LGR.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0004356; F:glutamate-ammonia ligase activity; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0009399; P:nitrogen fixation; IEA.
DR   InterPro; IPR004809; GlnA.
DR   InterPro; IPR001637; GlnA_adenyltn_S.
DR   InterPro; IPR008147; Gln_synt_beta.
DR   InterPro; IPR008146; Gln_synt_C.
DR   Pfam; PF00120; gln-synt; 1.
DR   Pfam; PF03951; gln-synt_N; 1.
DR   ProDom; PD001057; Gln_synt_C; 1.
DR   TIGRFAMs; TIGR00653; GlnA; 1.
DR   PROSITE; PS00180; GLNA_1; 1.
DR   PROSITE; PS00182; GLNA_ADENYLATION; 1.
DR   PROSITE; PS00181; GLNA_ATP; 1.
KW   Ligase.
SQ   SEQUENCE   469 AA;  51791 MW;  8FDA038A6EAAC14A CRC64;

  Query Match             53.1%;  Score 51;  DB 2;  Length 469;
  Best Local Similarity   52.9%;  Pred. No. 22;
  Matches    9;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GINWNGGSTGYADSVKG 17
              | :: ||:||:  ||||
Db        162 GTSYEGGNTGHRPSVKG 178


RESULT 11
Q7UUV9
ID   Q7UUV9      PRELIMINARY;      PRT;   838 AA.
AC   Q7UUV9;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   RB3034.
OS   Rhodopirellula baltica.
OC   Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC   Planctomycetaceae; Pirellula.
OX   NCBI_TaxID=117;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1;
RX   MEDLINE=22735913; PubMed=12835416;
RA   Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA   Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA   Schlesner H., Amann R., Reinhardt R.;
RT   "Complete genome sequence of the marine planctomycete Pirellula sp.
RT   strain 1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR   EMBL; BX294138; CAD72968.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   838 AA;  94147 MW;  3321CFAA9E4A0336 CRC64;

  Query Match             53.1%;  Score 51;  DB 16;  Length 838;
  Best Local Similarity   60.0%;  Pred. No. 42;
  Matches    9;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 INWNGGSTGYADSVK 16
              | | ||:  || |||
Db        218 IKWQGGTVAYATSVK 232


RESULT 12
Q9U234
ID   Q9U234      PRELIMINARY;      PRT;   767 AA.
AC   Q9U234;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Y56A3A.30 protein.
GN   Y56A3A.30.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Matthews L.;
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; AL132860; CAB60506.1; -.
DR   WormPep; Y56A3A.30; CE22595.
SQ   SEQUENCE   767 AA;  84076 MW;  B8D9E3C32EE8CA08 CRC64;

  Query Match             51.0%;  Score 49;  DB 5;  Length 767;
  Best Local Similarity   47.1%;  Pred. No. 77;
  Matches    8;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy          1 GINWNGGSTGYADSVKG 17
              |  : ||:||| |:: |
Db        718 GYGYGGGTTGYMDNLVG 734


RESULT 13
Q9WZJ8
ID   Q9WZJ8      PRELIMINARY;      PRT;   279 AA.
AC   Q9WZJ8;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein TM0739.
GN   TM0739.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
DR   EMBL; AE001744; AAD35820.1; -.
DR   PIR; F72339; F72339.
DR   TIGR; TM0739; -.
DR   GO; GO:0003824; F:catalytic activity; IEA.
DR   InterPro; IPR006674; HD.
DR   InterPro; IPR003607; Met_phsphohydro.
DR   InterPro; IPR006675; Unchar_HDIG.
DR   Pfam; PF01966; HD; 1.
DR   SMART; SM00471; HDc; 1.
DR   TIGRFAMs; TIGR00277; HDIG; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   279 AA;  32034 MW;  B1C6494321C85B3B CRC64;

  Query Match             50.0%;  Score 48;  DB 16;  Length 279;
  Best Local Similarity   60.0%;  Pred. No. 34;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 NWNGGSTGYADSVKG 17
              ||:|  ||| | :||
Db        200 NWDG--TGYPDGLKG 212


RESULT 14
Q9XGB4
ID   Q9XGB4      PRELIMINARY;      PRT;   298 AA.
AC   Q9XGB4;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chitinase.
OS   Trifolium repens (Creeping white clover).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Trifolium.
OX   NCBI_TaxID=3899;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Dutch white; TISSUE=Root;
RA   Crockard M.A., Broderick K., Pulvirenti G., Cooper J.E.;
RT   "Expression of a class 3 chitinase in white clover roots following
RT   homologous challenge.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ011940; CAB43737.2; -.
DR   HSSP; P23472; 2HVM.
DR   GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18AS.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   298 AA;  31340 MW;  A82451AC8E7885A4 CRC64;

  Query Match             50.0%;  Score 48;  DB 10;  Length 298;
  Best Local Similarity   50.0%;  Pred. No. 37;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 WNGGSTGYADSVKG 17
              :| | :||:|::||
Db        283 FNDGQSGGYSDAIKG 296


RESULT 15
Q9NPP6
ID   Q9NPP6      PRELIMINARY;      PRT;   416 AA.
AC   Q9NPP6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Immunoglobulin heavy chain variant (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pluvinet R., Estivill X., Escarceller M., Sumoy L.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Auffray C., Ansorge W., Ballabio A., Estivill X., Gibson K.,
RA   Lehrach H., Poustka A., Lundeberg J.;
RT   "The European IMAGE consortium for integrated Molecular analysis of
RT   human gene transcripts.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL389978; CAB97534.1; -.
DR   HSSP; P01789; 1MCP.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 3.
DR   PROSITE; PS00290; IG_MHC; 2.
FT   NON_TER       1       1
SQ   SEQUENCE   416 AA;  44786 MW;  8C41708BB8AB4687 CRC64;

  Query Match             50.0%;  Score 48;  DB 4;  Length 416;
  Best Local Similarity   43.5%;  Pred. No. 54;
  Matches   10;  Conservative    3;  Mismatches    4;  Indels    6;  Gaps    1;

Qy          1 GINW------NGGSTGYADSVKG 17
              |: |      :| :  ||||||
Db          4 GLEWVSRISSSGDTVDYADSVKG 26


Search completed: September 18, 2004, 03:46:29
Job time : 91.6429 secs
```

This Page Blank (uspto)